

THE ROLES OF ARTIFICIAL INTELLIGENCE IN EDUCATION: CURRENT PROGRESS AND FUTURE PROSPECTS

• David McArthur • Matthew Lewis • Miriam Bishay

RAND 1700 Main St. Santa Monica, CA 90407-2138

Summary

This report begins by summarizing current applications of ideas from artificial intelligence (AI) to education. It then uses that summary to project various future applications of AI -- and advanced technology in general -- to education, as well as highlighting problems that will confront the wide-scale implementation of these technologies in the classroom.

The earliest applications of AI in education developed intelligent tutoring systems (ITS). For the most part, ITS, like CAI systems before them, have attempted to implement traditional methods of learning and teaching. Drill-and-practice, and other variants in which students solve relatively short problems chosen by the teacher, have a proven pedigree in the classroom. Perhaps more important, ITS using a drill-and-practice method have generally aimed at well-defined and well-accepted goals for learning. These typically include factual knowledge and procedural skills, like algebra symbol manipulation, that are part of traditional school curricula, and that can be measured by existing standardized tests. Working with traditional methods of teaching and learning, and using traditional means of evaluating outcomes, the developers of ITS have tried to show that ITS can significantly improve the speed and quality of students learning. And, to some extent, they have been successful.

However a looming paradox undermines these successes. The technologies that make it possible to automate traditional methods of teaching and learning are also helping to create new methods, and to redefine valued educational goals. For example, new technologies can automate symbol-manipulation algebra and spelling correction, making these skills less important to learn, while increasing the importance of "higher-order" skills required to do creative mathematics and writing. As a result, attempts to use new technologies in education to further traditional learning goals or traditional methods of teaching makes less and less sense.

This situation poses tremendous difficulties to the development of effective educational technology applications. Traditional goals and methods for learning are at least well understood and relatively well-defined. But new methods -- for example learning through inquiry, collaboration, or visualization -- and new goals for learning have not yet been agreed upon by the educational community at large, let alone fully operationalized. Crafting effective technologies for fixed learning goals and methods of teaching and learning is challenging enough. But is even more demanding when the target you are aiming at is moving.

Microworlds and interactive learning environments (ILEs) are one response to these changes in the educational landscape. Generally, they are trying to implement an inquiry-based method of teaching and learning, perhaps helping to bring this method into the classroom on a large scale for the first time. However, at the same time, ILEs are involved in a shift of educational goals as well as methods. Its not simply a matter of providing tools for new methods of learning skills now taught in many classrooms. In some cases, new educational goals focus on topics that are not part of traditional curricula, such as boolean networks and chaos, graph theory, and inquiry skills themselves. In other cases the focus is on traditional topics, like fractions or polygons, but the intent is to foster a deeper conceptual understanding of ideas that are usually taught as simple procedures.

The movement from ITS to ILEs, and to mixed-initiative systems that represent a combination of both approaches, illustrates a general pattern in educational technology today. Virtually all important computer-based applications to education are not simply trying to teach traditional skills more quickly, efficiently or less expensively. Rather, like ILEs, they are participating in an attempt to change methods of learning and teaching and to redefine valued educational goals and learning outcomes.

As a result, there is no reason to believe that the most effective uses of AI (or advanced technology in general) in education will happen quickly or without careful policy and planning. In the short-term, technologies resembling many ITS -- that aim at well-defined learning goals and that can be moved into classrooms with a minimum of disruption -- will provide the most statistically significant improvements in student outcomes. Policies that support research based on their ability to generate such results in "horse race" evaluations risk encouraging technology applications that miss longer-term benefits.

On the other hand, policies which simply give researchers free reign to develop software that focuses on new methods of teaching and new learning outcomes also run considerable risks. The problems in developing these systems and moving them into education on a broad scale are not simply technical ones. Our experience provides a case in point. When developing ITS for mathematics education, 75% our effort was spent on technical and research issues, but implementation required most of our time when we integrated ILEs into schools. As technology continues to transform the goals for student learning and to enlarge the range of methods for teaching and learning, implementation will require proportionally more effort.

In general, implementation tasks must develop new curricula, assessment methods and instruments, teaching practices and professional standards, and teacher education. These tasks cannot all be effectively done by any single group of researchers. Rather, we believe that scaling up new technologies like ILEs will require a division of labor -- different groups or projects working in a coordinated fashion to put together the technology, curricula, assessment tools, professional standards, and teacher training pieces of a package of broad educational reform. However, today we see very little evidence of such coordination. Research in these areas is typically funded by disjoint government programs that provide few incentives or mechanisms necessary to engender needed cooperation. In the future it will be important to consider new policies that improve the communication between these groups. A variety of policy options are worth examining, including larger consortium-based projects that include expertise in software development as well as teacher training and other key

stakeholder groups, smaller separate projects that work together synergistically using new networking technologies, and incentives that bring high-tech companies into better cooperation with educational technology research and classroom practice.

Acknowledgements

This research was funded in part by the National Science Foundation (Applications of Advanced Technologies Program) under grants MDR-8751515 and MDR-9055573. Additional funding was provided by RAND. Views or conclusions expressed herein do not necessarily represent the policies or opinions of the sponsors.

Introduction and Overview

Computer technologies are changing the practice of research and business, and -- very slowly -- the content and practice of education are beginning to follow suit. This paper discusses how work in artificial intelligence (AI) is contributing new approaches to education and learning. The hallmark of AI applications in education is that they attempt to explicitly represent some of the reasoning skills and knowledge of expert practitioners, and to exploit that expertise for teaching and learning. In business we see growing evidence that information technologies are leading to substantial improvements in productivity by automating routine activities (Zuboff, 1988). Similarly, it seems that if we can impart basic cognitive skills of teachers to computers we might delegate some teaching to machines and thus improve educational outcomes. However, this paper will chronicle several ways in which such a theory vastly oversimplifies practice.

The main goal of this paper is to review past and present trends in the applications of AI in education, and to project from the successes and failures of those applications to possible future applications. The paper is a personal perspective in several senses. It selectively highlights a few important issues but makes no attempt to survey all key ideas in educational technology. Discussion is limited to two main applications of AI in education: intelligent tutoring systems and microworlds. Even more narrowly, the paper will compare these two approaches primarily in terms of their methods of learning and teaching -- the procedures, principles, and techniques they embody to facilitate learning -- and their learning outcomes or goals for learning

-- the particular kinds of knowledge and learning they help students acquire, or try to help them acquire. The paper is also a personal perspective because it illustrates many of its main points with reference to our own research at RAND over the past several years. We also discuss other research to support general points, but many outstanding systems and studies are mentioned only in passing or not at all.

The organization of the paper is roughly chronological. The first section discusses the application of AI ideas in developing early intelligent tutoring systems, reviewing both their successes and limitations. These limitations motivate the relatively recent development of interactive learning environments and microworlds, which are discussed in the next main section. The paper concludes with some general reflections on the current state-of-the-art, and with some speculations on possible future directions for AI in education.

Intelligent tutoring systems

This section discusses the development of intelligent tutoring systems for education over the last decade. We begin by reviewing the fundamental ideas behind these systems and illustrating them with reference to tutors we have developed. While knowledgeable readers will find this review familiar, it sets the stage for the following discussion of successes and failures.

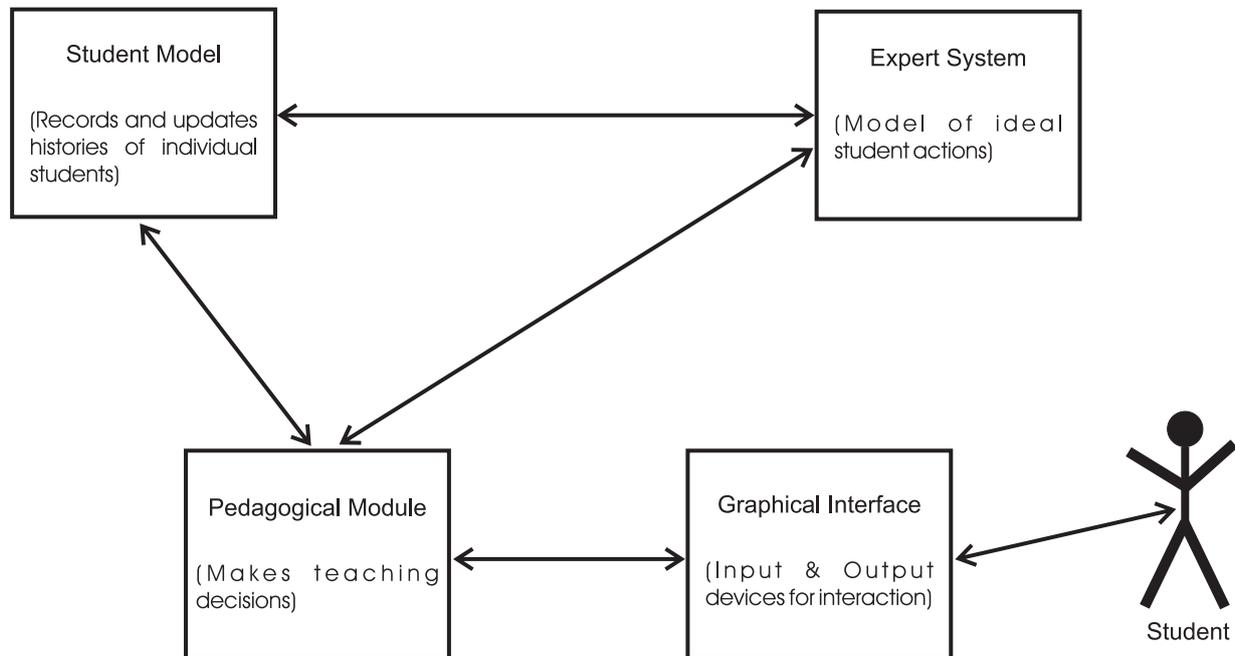
Anatomy of an Intelligent Tutoring System

The first (and still foremost) application of AI to education has been to build intelligent tutoring systems (ITS). (NOTE: In this paper, I will use "ITS" as an abbreviation both for "intelligent tutoring system" and the plural, "intelligent tutoring systems". "ITSs" just looks too strange.) ITS have been under development at least since WEST (Burton and Brown, 1982) and SOPHIE (Brown, Burton and deKleer, 1982) nearly twenty years ago. Many specific systems, their structure and goals, have been discussed in depth by Sleeman and Brown (1982), Wenger (1987), and Psofka, Massey, and Mutter (1988), among others. Ohlsson (1986) and Schank and Edelson (1990) have also contributed valuable critical reviews of the direction of the field. Here we review only a few key features of ITS.

ITS attempt to capture a method of teaching and learning exemplified by a one-on-one human tutoring interaction. For researchers in AI this method of teaching was a natural one to target first for several reasons. Drill-and-practice versions of one-on-one tutoring are relatively well-understood ways of communicating knowledge. This method of learning and teaching is widely accepted both by the educational community and by our culture as a whole. And it has achieved broad popularity for good reason. One-on-one tutoring allows learning to be highly individualized, and consistently yields better outcomes than other methods of teaching (Bloom, 1984). Although many methods have been examined, no other has reliably yielded "2 sigma" improvements -- usually over a full letter-grade -- in student outcomes. Failure to attribute strong outcome improvements to other methods of teaching and learning may be partly a function of inadequate techniques for evaluating novel learning outcomes, as we elaborate below. However, regardless of evaluation problems, it is clear that one-on-one tutoring remains a "gold standard" of learning.

Although ITS differ in a variety of ways, most have a characteristic structure. Figure 1 sketches a generic ITS. Superficially, ITS differ little from the CAI systems preceded them. In general, both are characterized by a common philosophy that includes high tutor control, and short-answer task format. In most CAI systems and ITS students learn by working a series of relatively brief questions. In both cases the system plays the exclusive role of the task expert, controlling the selection of tasks or problems, while the student responsible for answering them. The system also plays the role of the critic, and in most ITS the system rather than the student decides when critical feedback will be supplied. The main differences between ITS and earlier CAI systems do not reflect differences in methods of teaching and underpinning philosophies of learning. Rather, they reflect engineering and psychological enhancements that permit ITS to tutor in a knowledge-based fashion. Unlike previous CAI systems, ITS represent some of the knowledge and reasoning of good one-on-one human tutors, and consequently can coach in a much more detailed way than CAI systems.

ANATOMY OF AN INTELLIGENT TUTORING SYSTEM



(Adapted from Dede, 1992)

Figure 1: Components of a generic intelligent tutoring system

The heart of an ITS is its expert system. The expert system embeds sufficient knowledge of a particular topic area to provide "ideal" answers to questions, correct not only in the final result but in each of the smaller intermediate reasoning steps. The expert system thus allows the ITS to demonstrate or model a correct way of solving the problem. Often, like a human tutor, it can generate many different answer paths or goal structures (McArthur, Stasz, Hotta, Peter, and Burdorf, 1988). The same detailed data structures that expert systems generate in modeling expert reasoning also permit ITS to explain their reasoning at arbitrarily detailed levels. For example, if a student needs an explanation of why or how an algebra ITS did a step in solving an equation, the system might first say that it used the distributive rule. If the student requested more justification, it could elaborate by describing the terms that were distributed and the arithmetic "cleanup" steps that followed. Explanations thus turn expert systems from opaque "black box" experts into inspectable "glass boxes" (Foss, 1987).

ITS coach as well as model expert problem-solving. In particular, they can monitor the student as he or she solves a problem and can determine if every step is right. Thus, while questions were the atomic unit of discourse in CAI systems, in ITS the basic unit is the individual reasoning step. To support this detailed coaching, ITS often create and update a student model (see, among others Anderson, Boyle, and Yost 1985; London and Clancey, 1982; Sleeman and Smith, 1981). The student model reflects the correct rules the ITS thinks the student knows -- ones that are also found in the expert system or in the "ideal" student model (Anderson, Boyle, and Reiser, 1985). But, since most students are not ideal, the student model often contains rules that are "buggy"--invalid rules the ITS thinks the student believes (Matz, 1982; Sleeman, 1982). The ITS watches each step in the student's reasoning as he or she solves a problem. Each time the student makes an error, the ITS will diagnose the problem -- possibly updating the student model -- then attempt to remediate it with very detailed advice about how the expert system would do the step. This process repeats at every step in the evolution of a complete solution to a problem.

One part of most ITS that receives relatively little mention is the pedagogical component. While the expert system contains rules and knowledge that drive an ITS subject-specific performance, the pedagogical component is supposed to contain similar rules that encode expertise about tutoring itself -- for example when to interrupt students and what kinds of information to provide them. But, as we elaborate below, current ITS have little pedagogical expertise.

An Intelligent Tutoring System for Basic Algebra

The algebra ITS we have developed over the past several years illustrates many of the key features of intelligent tutoring systems (McArthur, Stasz, Hotta, Peter, and Burdorf, 1988; McArthur, Lewis, Ormseth, Robyn, Stasz, and Voreck, 1989). Our algebra ITS helps students learn freshman level algebra, focusing on equation solving and symbol manipulation. In different versions of the tutor the student can solve problems using symbols, operations, or commands. In the "Symbols" version, students input new equations that lead toward a solution. Students can input

equations either by typing or by using electronic pencil. In "Operations" the student issues requests like "add 30 to both sides of the equation". And in "Commands" the student specify very high level goals, like "collect" like terms. The different versions allow the student to focus on different levels in the rather complex hierarchy of reasoning skills that characterize even simple algebra. Students practice one level of decision-making, and the ITS takes care of other levels that the student is not concentrating on at a given time. For example, if the student says "collect" the ITS will figure out what "collect" means in terms of operations, then also do the appropriate symbol manipulation.

In Figure 2 the student is using the operations version of the algebra tutor. The operations the student chooses are recorded in the right side window. The left side contains various menus that allow the student and ITS to converse. For example the student can move scroll display, can ask for help, can ask the ITS to do a step and to explain what it was doing, and can create their own problem or get easier or harder ones.

Legend:

simplify	Reduce terms to a simpler form, or do arithmetic
/ ?	Devide both sides by ?
* ?	Multiply both sides by ?
+ ?	Add ? to both sides
- ?	Subtract ? from both sides
distribute ?	Expand a * or / using the distributive rule

Left Sidebar:

- INEQUALITIES
- Scroll Right, Scroll Left, Scroll Down, Scroll Up
- My Answer Ok?, Help Next Step, Explain Your Step
- Login, Erase Input, Move Box
- Homework Problem, Student Problem, Easier Problem, Harder Problem
- QUIT

Central Workspace:

$$-2 + 7x = 9(-6 + x) - 4(-6 + x)$$

$$-2 + 7x = -54 + 9x - 4(-6 + x) \quad -2 + 7x = 5(-6 + x)$$

$$-2 + 7x = -54 + 9x - 4(-6 + x) \quad -2 + 7x = -30 + 5x$$

$$-2 + 7x = -54 + 9x - 24 + 4x \quad -2 + 7x + 30 = 5x$$

$$-2 + 7x = 5x - 30 \quad 28 + 7x = 5x$$

$$-2 + 7x + 30 = 5x \quad 28 + 7x = 5x - 7x$$

$$28 + 7x = 5x \quad 28 = 5x - 7x$$

$$28 = 5x - 7x \quad 28 = -2x$$

$$28 = -2x \quad 28 - 5x = -7x$$

$$\frac{-28}{2} = x \quad \frac{-28}{2} = x$$

$$-14 = x$$

Right Sidebar (Operations Log):

- (collect) $9(-6 + x) - 4(-6 + x)$
- (distribute) $-6 + x$
- (+ 30)
- (simplify)
- (- 7x)
- (collect x)
- (/ -2)
- (Simplify)

Bottom Section:

CONDITIONS: There is more than one x on one side of $-2 + 7x = 9(-6 + x) - 4(-6 + x)$.

GOAL: A useful goal was to COLLECT the variables together.

OPERATION: To carry out this goal the variables $9(-6 + x) - 4(-6 + x)$ Were combined to make $(9 - 4)(-6 + x)$.

Figure 2: An intelligent tutor for basic algebra

When the student creates steps in a solution he or she can develop multiple different solution paths that appear as different branches in a solution tree. Solution trees "reify" the student's reasoning process by showing connections between steps (Collins and Brown, 1987). Such reification allows the student to easily compare different solution approaches, or to contrast a "buggy" approach with a correct one. In this case, for example, the student's first (leftmost) attempt began by distributing and the second attempt uses the more elegant approach, by first collecting. This approach was suggested by the computer after a request for help from the student (the command that began this branch is in inverse video in the commands window to the right, indicating this step was done by the machine). At the point shown in the Figure, the student has asked the ITS to explain how and why it did the step. Explanations appear in the bottom window. The student can request explanations at increasing levels of detail. Like other ITS, the algebra tutor's main strength is that it permits coaching at an arbitrarily fine grain size. As Anderson has pointed out (Ohlsson, 1986), it would be difficult -- perhaps impossible -- for a human tutor to provide any more detailed feedback on the logic of algebraic symbol manipulation.

We have experimented with two very different pedagogical components in our algebra tutor. In the version we have been discussing, the pedagogical policy permits high students control. The student decides when to ask for help, when to request the expert to do and explain, when questions should be tougher and easier. Another version of the system is at the other end of the pedagogical spectrum. It is completely tutor controlled; the tutor, not the student, decides what questions to give, and when to help the student.

Successes of Intelligent Tutoring Systems

Most ITS are still being used on a very small scale, and only a few have been tested widely. Of these, just a handful claim to improve students' outcomes in the classroom using standardized tests. Successful ITS have been mainly restricted to areas of mathematics and science, where it is both easier to build ITS and easier to measure learning outcome improvements. Perhaps the most thoroughly tested ITS are Anderson's Geometry and Lisp tutors

(Anderson, Boyle, and Yost, 1985; Anderson and Skwarecki, 1986; Schofield, Evans-Rhodes and Huber, 1990). Anderson claims "2 sigma" improvements using his Geometry and Lisp tutors; thus, for these limited topics, he may be approaching Bloom's gold standard. SHERLOCK, a tutor for electronic troubleshooting, has resulted in similarly dramatic performance improvements (Lesgold, Lajoie, Bunzo and Eggan, 1993; Lesgold, Eggan, Katz, and Rao, in press).

Our algebra ITS has been tested on several occasions using introductory algebra classes at a local high school (McArthur and Stasz, 1990; Robyn, Stasz, Ormseth and McArthur, 1989). In each fielding, one or more classes used the ITS and comparison classes learning the same curriculum, did not. Our evaluations have primarily focussed on the role of the tutor in improving traditional algebra learning outcomes. Specifically, we were interested in measuring improvements in solving symbolic equations, and on seeing whether students could benefit by reasoning separately about goals, operations, and symbols. Our best results indicate that students using the algebra tutor scored almost a letter-grade higher than comparison students (see Robyn, Stasz, McArthur, Ormseth and Lewis (1992) for a more detailed discussion of curriculum and results).

However, these kinds of narrow results do not tell the whole story. In addition to enhancing students' outcomes in traditional algebra symbol manipulation, several evaluations of our algebra tutor have attempted to see if the tutor could improve skills not measured by most standardized tests. For example, might students using the algebra ITS improve their ability to model realistic situations with equations, as well as skills for manipulating and solving equations? Could they improve other "higher-order" skills, like planning, debugging (fixing errorful solutions) and goal setting? As we report in more detail elsewhere (Robyn, Stasz, McArthur, Ormseth and Lewis, 1992), there is qualified evidence that our algebra ITS can lead to such outcomes. However, these benefits come at a cost. To improve these skills we had to extensively restructure the standard algebra curricula in which our ITS was embedded. For these pilot tests, crafting the software took less time than developing the non-computer materials and providing teacher training.

We see in these results an important general pattern. By adopting one-on-one tutoring, ITS have aspired to implement a very familiar, well-understood, and accepted method of teaching and learning. At the same time, most ITS have aimed at equally familiar learning goals or outcomes. These typically comprise procedural skills, like algebra symbol manipulation, that are part of traditional school curricula. Such outcomes are relatively easily measured by the standardized tests that are used in most school curricula. Consequently, it is relatively straightforward to determine the success or failure of ITS in objective and broadly accepted ways. On the other hand, to the extent that ITS -- or any technology -- attempts to target more novel learning outcomes, it becomes both more difficult to measure success, and more demanding to implement the technology in traditional classroom settings.

Reasons for ITS successes

There has been surprisingly little reflection in the literature on the reasons why ITS have succeeded, and on why their successes have been limited. This kind causal analysis is important for several reasons. First, as Schofield's (Schofield, Evans-Rhodes and Huber, 1990) analysis of Anderson's Geometry tutor underscores, the actual reasons for success may not be the ones initially expected. The ability of an ITS, for example, to finely tune instruction may be less important than the fact that it permits teachers to spend more time with slower students or that it increases student motivation. Second, a given ITS, or any other demonstration system, has limited value if not tied to some general observations or principles. Ohlsson (1991) has persuasively argued that the real importance of any tutoring system is a function of the principles of learning which it illuminates. Any system, viewed as a software product, is likely to have a brief life-span and will directly impact few learners. However, principles of learning, validated through demonstration systems, add to the enduring knowledge of learning and teaching that can guide the construction of new generations of systems. They become part of an incrementally growing base of scientific and design knowledge. We disagree with Ohlsson's view that the only important principles on which ITS can impact are those that underpin cognitive theories

of learning. For one thing, teaching principles and general interface design ideas are equally important, and may not be reduceable in any systematic way to learning principles. But we agree that a key contribution of ITS lie in the general truths they uncover or exemplify. From this perspective, then, we believe several features or principles are responsible for the (admittedly limited) success that ITS now enjoy.

Micro-tutoring.

Probably the greatest strength of ITS are their ability to generate highly detailed feedback about problem solving -- to micro-tutor. They can coach and model problem solving down to "atomic" levels of reasoning. Micro-tutoring is the main teaching principle ITS embody that distinguish them from earlier generations of CAI. The converse learning principle argues that learners need rich, variable granularity feedback (Anderson, Boyle, Farrell and Reiser, 1987). When learners accomplish a task they use their skills along with external tools to generate reasoning and visible performance. Although learning can happen with nominal feedback, generally richer feedback yields more accurate diagnosis of errors, thus faster learning. The need for rich feedback is especially important when tasks are authentic and skills are embedded. Because so many skills may be used in the process of generating an answer or step, it may be difficult for students to locate their errors among many acceptable actions (often referred to as the "credit assignment" problem) and to draw a general inference from the errors (the "repair" problem), without detailed feedback.

Tutor control of learning.

Although some ITS, such the student-controlled version of our algebra tutor, permit limited student choice, for the most part interactions with ITS are tightly controlled by the software. In most cases, the ITS selects the next task or problem, decides when the student needs support and feedback in problem solving, and determines the nature of the information the students receive. Students may tailor information; for example they may request more detailed explanations. But their latitude is usually highly circumscribed. The principle of high tutor control reflects an implicit belief that a competent tutor is usually in a better

position to make decisions about what experiences and information students need to learn effectively than the students themselves. Of course, this assumes, at a minimum, that the tutor knows the content the students want to learn, and also knows the students' specific knowledge state -- what they know, and what knowledge they lack -- at any given time. The expert systems and student models of ITS attempt to provide this expertise and to thus meet the demands of high tutor control of learning.

Impasse-driven coaching.

A related feature of ITS is that they are stimulated to action by student difficulties or impasses. In the tutor-controlled version of our algebra tutor, for instance, feedback and help is triggered by an algebraic error. By organizing learning around small short-answer tasks, and by choosing tasks that are demanding for students, ITS attempt to maximize opportunities for impasses. But, like many ITS, our tutor's decision-making in response to impasses is quite "thin". Interventions are immediate reactions to errors alone. For example, they are not conditioned by a plan to set up the learning experience beforehand or to help students define goals or plans. However, whether ITS reason extensively or not, virtually all reasoning attempts to recognize student impasses and overcome them. The learning principle that corresponds to teaching principle of impasse-driven coaching is to provide immediate feedback (Anderson, Boyle, Farrell and Reiser, 1987). Most ITS, implicitly or explicitly, are built on the premise that a good learning system will provide detailed feedback as soon as an impasse is detected.

ITS fit well into classrooms.

To the extent that ITS appear to contribute to effective learning outcomes, they substantiate the principles of micro-tutoring, high tutor control, impasse-driven coaching and providing rich and immediate feedback. However, although ITS are successful in part because they are consistent with various theoretical principles of learning and teaching, practical reasons may be equally important, if not more so. The simple fact is that ITS actually fit quite well into existing classrooms, easily filling the shoes of earlier CAI programs and integrated learning systems that have enjoyed at least modest success.

ITS are congruent with existing classroom practice in at least two senses. First, they generally aim at learning goals or outcomes that are already embedded in traditional curricula -- algebra symbol manipulation, programming, and geometry, for example. Second, they adopt a popular method of teaching and learning. Most classrooms still combine lecture with drill-and-practice. As a consequence, teachers have little trouble finding an effective role for ITS like our algebra tutor. They can usually be plugged into existing curricula with minimal change to course plans; for example, they often simply replace pencil-and-paper homework or drill-and-practice seatwork.

It is important to note that not all software fits into classrooms so easily. As we elaborate below, many recently developed computer environments are more aligned than ITS with emerging standards for mathematics curricula which advocate new methods of teaching and learning and emphasize new goals for student learning (e.g., National Council of Teachers of Mathematics, 1989). Yet, regardless of their popularity in research, these environments do not always work well in traditional classrooms. Later sections will have more to say about the demands of implementing various kinds of systems in the classroom.

Limitations of Intelligent Tutoring Systems

While ITS have been somewhat successful on a small scale, several problems must be overcome before they have widespread impact. Various authors (e.g., Wenger, 1987; Psotka, Massey, and Mutter, 1988) have discussed a wide range of limitations. Many of these challenges can be predictably factored by ITS component -- limitations associated with the expert system, student model, pedagogical component, and interface. In this section we touch relatively briefly on just a few shortcomings that we believe are most fundamental and which may not be overcome simply through incremental improvements to various ITS components. In the next section we discuss how researchers applying AI to education are responding to these challenges.

Limitations in learning outcomes

Educational technologies can aim at a wide variety of

learning goals or outcomes, from helping students learn skills in traditional subjects and curricula, to making new topics accessible to younger students, to facilitating deep conceptual understanding, to fostering metacognitive skills like debugging. Most ITS have focused on subjects taught in typical primary- and high-school level courses. In this context, probably the most significant limitation of ITS is that, to date, they can be developed only for a few topic areas. Effective ITS require virtual omniscience -- a relatively complete mastery of the subject area they are to tutor, including an understanding of likely student misconceptions. Thus, the most successful ITS have been developed for simple "closed worlds" and procedural skills like solving short problems in mathematics, science, and logic. These are the easiest topics for which to build complete expert systems simply because cognitive science provides elaborate task analyses of competence in these areas. Such complete knowledge of a subject, and possible errors in knowledge, is also important for developing detailed student models.

Conversely, ITS have not yet been successfully developed for less well-understood or well-defined subjects. These include a wide range of topics in history and social sciences, where natural language understanding would appear to be prerequisite for any effective ITS. But they also include areas in science and mathematics, such as those emphasized by NCTM (National Council of Teachers of Mathematics, 1989). For example, cognitive science cannot yet provide a complete task analysis of expertise in posing novel problems (Brown and Walter, 1990), designing good experiments, or creatively applying the scientific method. Similarly, while ITS can help students practice the symbolic manipulation of equations, they cannot tutor the effective application of these representations in modeling real-world situations. As a result, it is currently impossible to develop an omniscient expert system for these topic areas, let alone modeling student misconceptions. And, without an effective expert system, an ITS loses its educational leverage.

Subject competence limitations of ITS can be interpreted weakly or strongly. A weak interpretation agrees that effective ITS are currently limited to topic areas which are relatively simple, and perhaps of minor importance in

emerging curricula. But a weak interpretation also argues that eventually ITS will be built for many of these more interesting topics, perhaps relying on new breakthroughs in the representation of knowledge provided by artificial intelligence and cognitive science research. On this interpretation, the subject limitations of ITS are little more than a reflection of the youth of the field.

On the other hand, a strong interpretation of the competence limits of ITS argues that it may never be possible to develop ITS for interesting topic areas. One of the reasons that rote calculation and algebraic symbol manipulation are no longer highly valued parts of math curricula, for example, is that these computations can be automated. Since they can be done by machine, so the argument goes, they lose value as curriculum topics, while new skills (that perhaps presuppose these simpler ones) become increasingly important. In math, for example, schools are slowly reducing the time spent learning multiplication tables, while reluctantly increasing exposure to "higher-order" skills involved in finding and understanding abstract mathematical patterns (Steen, 1990).

The strong interpretation of the competence limits of ITS thus does not deny that new research in artificial intelligence will bring more topic areas in range for ITS. Rather it suggests that the very act of automating new kinds of reasoning will devalue them as curriculum topics. On this argument, the developers of ITS are caught in a "catch 22" they can never escape. As ITS improve, the learning outcomes they can engender will expand accordingly; but the value of these outcomes will decline at the same time.

Limited teaching and pedagogical expertise

A second important limitation of ITS pertains to their restricted knowledge of teaching. As we noted, most ITS have very impoverished pedagogical components. Such components often comprise a collection of rules that just seem to work reasonably well in practice. There is no scientific encyclopedia of good tutoring heuristics to consult, let alone a principled theory of one-on-one tutoring from which specific heuristics might follow. Cognitive science has not yet progressed to the point where it can offer good task analyses of pedagogical

expertise the way it has articulated the reasoning in "closed worlds" like symbol manipulation algebra (Matz, 1982), arithmetic (Brown and Burton, 1978), and chess (Newell and Simon, 1972). In short, the pedagogical component of ITS is an "expert system" that we cannot yet build.

Just as learning outcome and subject limitations have weak and strong interpretations, pedagogical limitations can also be viewed narrowly or broadly. The narrow interpretation, implied above, argues that to improve the pedagogical capabilities of an ITS we need to enhance the rule-base that tells it when and how to coach students. On this interpretation, the basic structure of tutorial interaction -- a method of teaching that emphasizes one-on-one tutoring, and the more detailed principles of micro-tutoring and impasse-driven coaching -- remains intact. All that is needed is to enrich the pedagogical knowledge base. In short, a weak interpretation suggests that we simply need to do a better job of implementing the method of teaching and learning that ITS are already pursuing.

A strong interpretation of ITS pedagogical limitations argues the problem is not merely that the prevailing one-on-one tutoring method is incompletely developed. Rather, this interpretation suggests a more fundamental problem -- that ITS are constrained to a single method of teaching and learning while truly expert tutors can adopt different methods. However good ITS are at micro-tutoring, they are still limited to a drill-and-practice style of interaction. More generally, they lack the ability to tutor flexibly; to adopt different teaching methods when appropriate, and to permit students to use different learning styles. By contrast, competent one-on-one human tutors may shift methods depending on the needs of students and on other contextual factors. A session may begin in an apprenticeship mode, then shift into drill-and-practice, and finally into less constrained, student-centered inquiry. This "meta level" reasoning about appropriate teaching and learning methods is well beyond today's ITS.

The drill-and-practice method of teaching embedded in ITS appears to be more suitable for tuning existing knowledge than for conceptual learning of substantial pieces of new knowledge (Ohlsson, 1991). This limitation, in part, reflects limitations in the cognitive science that

underpins ITS. Psychological theory has accumulated a good understanding of how to remediate small "bugs" in knowledge. For example, BUGGY provides a detailed model of the microgenesis of misconceptions in arithmetic (Brown and Burton, 1978); and Anderson's ACT* shows how new knowledge can be built incrementally from existing rules through operations like composition, generalization, and specialization (Anderson, 1983). This research provides a solid foundation for ITS that help student make minor adjustments in knowledge. But cognitive science does not have a comparable understanding of the initial learning of large chunks of knowledge, and consequently, we lack a rigorous theoretical basis that would guide the design of learning systems for acquiring conceptual knowledge.

Current trends in the application of AI in Education

In the last several years, applications of AI in education have diversified; approaches are more fractured, and the field is certainly not as unified as five years ago when ITS dominated AI research in education. These new trends exemplify different responses to the difficulties encountered in developing and testing the first generation of ITS. Very broadly, while some recent research attempts to improve the one-on-one tutoring method of teaching associated with ITS, other work is investigating different methods of teaching and learning. This research is rethinking important principles of learning and teaching that should underpin education in general and the design of educational technology in particular. At the same time, new work is attempting to expand the range of learning goals and outcomes associated with AI-based systems for education. In some cases this means developing educational software for a more diverse set of subjects, but in other cases the targeted knowledge has less to do with the subjects learned than with the quality or depth of learning.

In this section we give an overview of several current trends. As in previous sections, our intent is not an exhaustive review of current research and results; instead, we are interested in characterizing the main thrusts.

Continued Development of ITS

A substantial amount of work continues within the ITS framework. Different groups are attempting to improve the

various components of ITS, to develop applications in increasingly complex subjects, and to make ITS more cost-effective to develop by providing "shells" and other system-building tools that institutionalize the basic structure of ITS. In doing so, this research largely holds fixed the drill-and-practice method of teaching, while attempting to enhance learning outcomes or goals either by improving the quality of tutoring of subjects already within range of ITS, or by expanding the range of subjects ITS can tutor.

New subject areas.

While many early ITS focused mainly on simple topics in high school mathematics, recently ITS have been developed for more advanced topics in mathematics (Du and McCalla, 1991) and science (Lester and Porter, 1991). ITS have also grown beyond mathematics and other more formal subjects to include topics in history, language and social science. Bruneau, Chambreuil, Chambreuil, Chanier, Dulin, Lotin and Nehemie (1991) describe the design of a tutor for reading; Frederiksen, Donin, DeCary and Edmond (1991) are developing a tutor for second language learning; and Feifer (1989) has developed a tutor that not only helps students learn to read but also focuses on inference and knowledge-structuring strategies. Similarly, ITS have diversified beyond public school curricula to topics in training and vocational education. For example, new ITS for electronics, maintenance, and troubleshooting (Cooper, 1991; Frederiksen, White, Collins and Eggan, 1988; and Kurland and Tenny, 1988) have built on the seminal work on SOPHIE (Brown, Burton and deKleer, 1982).

Enriched expert systems, student models, and pedagogy.

In addition to extending ITS to new subjects, ITS have also been enriched along several dimensions, improving the way their expert systems reason (e.g., Clancey, 1987), how they develop and use student models (e.g., Burns, Gray, and Radlinkski, 1991), as well as how they fashion tutorial interventions. For example, although pedagogical components remain impoverished, some ITS are now capable of relatively subtle reasoning when managing student impasses. In earlier ITS the response to student impasses was usually a simple function of the impasse itself; thus, a student making the same error ten times would

always receive the same feedback each time (Ohlsson, 1986). Now several ITS are capable of tutorial planning (MacMillian, Emme and Berkowitz, 1988; Woolf, 1991). These plans take several factors into account in generating feedback, in some cases including the students' past history of successes and errors. Planning is also beginning to play several other roles in ITS. In the past, virtually all the intelligence of an ITS was focussed on remediating the current impasse. Not only were ITS unable to take broader context into account in generating feedback, but they were equally unprepared to modify the sequence of tasks or problems given to the student in response to past performance. However recent planning techniques now permit some ITS to reason extensively about the features new tasks should have (McArthur, Stasz, Hotta, Peter and Burdorf 1988).

Improved interfaces, bandwidth, and visual representations.

Some researchers are attempting to improve the performance of ITS by keeping the same underlying ITS and exploiting better, more "user-friendly", graphical user interfaces (GUIs). Simpler communication between student and tutor, higher bandwidth dialogue, and visual explanations that are easy to understand as well as entertaining, may enhance learning outcomes substantially (Bonar, 1991). One example of an effective use of GUIs is GIL, a graphical version of Anderson's Lisp tutor (Reiser, Beekelaar, Tyle and Merrill, 1991). GIL has been experimentally compared to the original version of the tutor, which includes no graphical front end. The results showed, perhaps not surprisingly, that the graphical interface contributed significantly to improvements in student outcomes.

Indeed, in the short-term it is very likely that better interface design will contribute more to improved effectiveness of ITS than will enrichments in the expertise underlying their reasoning. This is not to say that knowledge is unimportant to tutoring. However, ITS have arisen from AI and cognitive science, and so they have focused primarily on the importance of teachers' knowledge of subjects and students in crafting systems. Effective interfaces may be as crucial for learning and teaching as high quality knowledge bases because they focus on other key

learning variables, including the roles of motivation, broad communication channels, and multiple different representations. Moreover, effective interfaces are easier to build than intelligent expert systems. As we elaborate in the next section, making ITS smarter is simply much tougher than making them clearer, more concrete, and more accessible.

Basic Research in Teaching, Learning, and Knowledge Representation

One reason that progress in developing ITS has slowed is that it depends on a groundwork of research in artificial intelligence and cognitive science. Cognitive science has supplied task analyses of various skills that permit ITS developers to implement detailed cognitive models. And artificial intelligence, primarily in the form of expert systems and production rule architectures, provides convenient vehicles for representing knowledge and the processes which apply it. As long as ITS could borrow from this underpinning research, progress in constructing ITS was relatively rapid. Development has been further facilitated as AI programmers (e.g., Forbus, 1991) and ITS researchers themselves (e.g., Anderson and Pelletier, 1991) create new languages, tools, and shells that simplify the construction of ITS expert systems and student models.

Cognitive science continues to provide better and more detailed cognitive models and task analyses for ever more sophisticated kinds of reasoning and problem solving. This research has advanced from beginnings where only well-defined and closed worlds like logic, puzzles, games, and algebra were understood (e.g., Newell and Simon, 1972). Today, for example, we have information processing models of problem solving in knowledge-rich subjects like medicine (e.g., Clancey, 1987), physics (e.g., Larkin, 1980), and electronics troubleshooting (e.g., Frederiksen, White, Collins and Eggan, 1988).

But this foundational research only extends so far. In spite of recent advances, many topic areas that educators regard as increasingly important still lie at least in part beyond current psychological theory. In fact, the gap between the knowledge and skills we understand at a detailed cognitive level and those we believe students should learn may be widening. For example, in mathematics curricula, simple

procedural skills to solve equations and more complex procedural skills of differential and integral calculus, used to occupy substantial roles in K12 curricula. However, as we noted, new standards have shifted emphasis to a somewhat ill-defined collection of "higher-order" problem solving skills. At present, cognitive science cannot offer rigorous task analyses of these skills. And without that research base, it is not possible to develop expert systems and student modeling facilities required by an effective ITS.

Having reached a point where ITS can no longer simply borrow from an existing foundation of research in cognitive science, educational demands are now becoming a stimulus for new basic research in cognitive science and artificial intelligence. Pedagogical expertise represents a good case in point. As mentioned, the pedagogical components of ITS are often impoverished, hampered by the absence of a rigorous task analysis of even rudimentary one-on-one tutoring methods. Over the past few years, this gap has stimulated studies that are now beginning to provide more detailed information processing models of human pedagogical expertise (e.g., Leinhardt, 1989; McArthur, Stasz and Zmuidzinas, 1990; Merrill, Reiser, Ranney and Tafton, 1992; Putnam, 1987). By expanding the cognitive science research base, this work may provide a new foundation for improved pedagogical components in the next generation of ITS.

Expert systems technology, borrowed from artificial intelligence, has similar limits. At least since Clancey's pioneering work using MYCIN as a foundation for tutoring in GUIDON (Clancey, 1987), it has been recognized that expert systems do not capture all the knowledge of an expert practitioner. Expert systems have been built for performance, not for teaching or explanation. For example, MYCIN can make medical inferences if it is supplied with the rules relevant to its narrow diagnostic task. All the deeper understanding of "first principles" and rich medical concepts can be compiled out of the knowledge-base with no loss of inference accuracy; in fact such compilation enhances inference speed dramatically. But, from an educational perspective, representing this deeper knowledge of first principles is more important than the surface rules. In diagnosis, as in any similarly complex cognitive skill, the ability to memorize

and apply a set of procedural rules is less educationally valuable than understanding the meaning and genesis of diagnostic decisions. In general, tutors built on the foundation of a traditional expert system tend possess relatively "thin" subject knowledge and are therefore capable of imparting only a shallow understanding of topics to students. This is as true of our algebra tutor as it was of Clancey's GUIDON. For example, students learning with our algebra tutor can expect to become proficient at manipulating symbolic equations, but they cannot expect to learn how equations can model real-world situations.

Just as various ITS shortcomings have stimulated new basic research in cognitive science, they have also begun to lead to improved representations of knowledge in artificial intelligence. Clancey (1987), for example, recognized that MYCIN's representations of knowledge needed to be fundamentally restructured to make them useful for effective tutoring. As Clancey explains:

"In building GUIDON we thought we were simply being "applications engineers" by making use of MYCIN's explanation facility in a tutorial setting. It was surprising ... to find out how little the original MYCIN explanation facility could accomplish for a student... The study of GUIDON's shortcomings has led to a radical change in our conception of MYCIN's rules, and supplies a new epistemological framework for building expert systems". (Clancey, 1987; pp. 168)

Research in qualitative physics (e.g., deKleer and Brown, 1984; Forbus, 1984) has also provided rich functional representations of processes and mechanisms. While much work in qualitative physics proceeded independent of ITS development, qualitative models were initially intended to provide rich explanations to learners -- a critical role in ITS as well.

New Methods of Learning and Teaching

As ITS have tackled new subjects and extended their capabilities in different ways, they have come to share less with the generic ITS sketched in Figure 1, and with one another. Some systems attempt to deepen subject understanding through new representations and expert systems, while others attempt to exploit effective interfaces

to overcome shallowness in knowledge and to weaken the assumption that an ITS must be virtually omnipotent. Nevertheless, ITS generally share a common view. They embody a one-on-one tutoring method of learning, underpinned by principles of high tutor control, impasse-driven coaching that is individualized to students' needs as the tutor sees them, and rich, fine-grained, immediate feedback.

Over the past few years a new collection of systems -- some embedding ideas from AI and some not -- have challenged this method of teaching and learning and the principles upon which it rests. These systems are even less of a kind than ITS, in part because they are not associated with a single competing method of teaching. Rather, collectively they are attempting to explore an increasingly wide range of methods of teaching and learning, and are pursuing a similarly broadened set of learning goals or outcomes.

One method of learning that has been most fully explored over the last five years we will refer to as inquiry-based, although it has also been described variously as student-centered, constructionist (Papert, 1980; Papert and Harel, 1991), constructivist (Davis, 1991), and discovery-based (Ausubel, 1961; Bruner, 1961). Systems that implement inquiry-based learning are structurally and conceptually much more diverse than ITS. For lack of a better name, we will refer to these systems collectively as interactive learning environments (ILEs). Although diverse, they share several principles that contrast them in fundamental ways to the views implicit in ITS. Broadly, the principles that tie ILEs together include:

- *Construction not instruction.* Students learn most effectively by constructing their own knowledge, not by being told through lecture, nor through organized drill-and-practice.
- *Student control not tutor control.* It follows that students must have significant, if not exclusive control of the learning interaction. At most, the tutor is viewed as a guide, not a god.
- *Individualization determined by the student not the tutor.* ILEs agree with ITS that individualized feedback and information are key to learning. However, they

differ on where personalized information should originate. While the tutor is responsible for tailoring feedback in ITS, in ILEs students generally receive unique feedback and information as a function of their interaction with the environment built into the system. Although the precise mechanisms of individualization differ substantially among ILEs, individualization is under at least partial control of the student.

- *Rich feedback generated by the student's interaction with the learning environment not the tutor.* Again, ILEs agree that fine-grained feedback is critical to learning. But they suggest that this rich information should arise as a function of the student's choices and actions in the learning environment, rather than as discourse generated by the tutor.

The principles that underpin an inquiry-based method of learning and teaching lead to designs for learning systems that differ substantially from ITS. Very broadly, the intelligence invested in ILEs is distributed across a range of tools rather than centralized in a tutor. These computer tools often include interactive video or other graphical representations, and they permit students investigate and learn topics largely free of external control. While this freedom derives from principled arguments about how effective learning happens, it also has practical benefits. ILEs are not as knowledge-intensive as ITS. They often benefit from some explicit representation of the topics that student's investigate, but they need not be omniscient; they do not (in some cases cannot) "know all the right answers". Further, since ILEs do not attempt to tutor, they are freed of obligations to model students' cognition and to make complex pedagogical decisions.

On the other hand, ILEs face their own challenges. If students are given "power tools" that magnify their ability to discover interesting ideas (Pea, 1987), what prevents them from using this power to flounder in a vast sea of uninteresting issues? How will they know what kinds of knowledge to construct? How will we judge such constructions? In the next section we discuss in detail one particular kind of ILE -- microworlds -- looking how they

confront some of the challenges that face ILEs. The final sections of the paper compare the approaches of ITS and ILEs, both looking at current accomplishments, and, more importantly, at future prospects.

Mathematical Microworlds

As in the previous section, we begin our discussion microworlds with a simple review of the basic concepts and examples from our research. This provides a foundation for an analysis of their successes and failures.

Methods of Learning and Learning Goals in Microworlds

Microworlds both move from tutors to tools and from a drill-and-practice method of teaching to an inquiry-based method of learning. Microworlds also represent a shift in the desired outcomes of learning. Many microworlds, including ours, have two distinct kinds of goals for student learning. First, as with ITS, it is usually important for students to learn subject-specific knowledge. More precisely, students often attempt to characterize the patterns of relationships among objects and properties that define the world. In SMITHTOWN (Shute and Glaser, 1990), for example, students might learn about the law of supply and demand by studying changes in costs of commodities as various factors influence supply or demand. Second, either implicitly or explicitly, many microworlds encourage students to learn inquiry skills themselves. These are relatively generic skills that students must know to conduct inquiries concerning virtually any topic. There is no rigorous task analysis of inquiry comparable to more "well-defined" skills, like factoring quadratics or performing an integration. However, following Lakatos (1976) and Steen (1988), we have compiled what we believe is a consensus view of inquiry activities, summarized in Table 1. These inquiry skills have been championed in the past (e.g., Polya 1962) and more recently (e.g., Schoenfeld, 1985; National Council of Teachers of Mathematics, 1989). Proponents see these inquiry skills as intrinsically valuable "higher-order" thinking abilities (Resnick, 1987). Thus, in microworlds inquiry is viewed both as a method of learning specific subject areas, and as a topic of learning.

RESEARCH PAPERS

Skill	Description
Proposing an issue.	An issue is not a specific problem that has an answer, or an hypothesis that is right or wrong. Rather, like Brown and Walter's (??) "Situation", an issue is a localized area of inquiry with features that can be taken as given, or challenged and modified. For example, in our <i>Polygons</i> microworld, an issue might be to look at the relationship of area and number of sides, as perimeter is held constant.
Proposing a specific problem.	A problem defines the specific focus of inquiry at a given time. For example, a problem might be to find a function that expresses area in terms of other polygon properties.
Generating conjectures or hypotheses.	In our microworlds, hypotheses can be quantitative or qualitative. In <i>Polygons</i> a qualitative hypothesis might be that area increases non-linearly as the number of sides goes up and perimeter is held constant.
Gathering observations that bear on an issue or hypothesis.	For example, generating a set of polygons by systematically increasing the number of sides while holding perimeter (or another scale parameter) fixed.
Confirming/disconfirming hypotheses.	Determining if the data collected is consistent or inconsistent with predictions derived from an hypothesis, or if it does not bear on the hypothesis at all.
Refining hypotheses.	This may mean repairing an hypothesis that has been disconfirmed, or forging a more detailed hypothesis from a vaguer, qualitative one. For example, once the student has confirmed a qualitative relationship between area and the number of sides, he or she may attempt to develop a specific function relating these properties.
Explaining or proving a hypothesis.	Decomposing a hypothesis into "smaller" hypotheses that are already proved. Formal proof is a special case of explanation. But informal explanations are also encouraged, as when a student relates the equation for interior angle of polygons to the simpler for exterior angle.

Table 1: Inquiry Skills

Overview of our Microworlds

The mathematical microworlds we are developing illustrate features that characterize many microworlds. Our microworlds are self-contained software environments in which students can create different kinds of mathematical objects (e.g., polygons) that have different properties or features (e.g., polygon properties include N , (the number of sides) $> a$ (apothem angle), and so on). The students then pursue inquiries to find and understand the patterns of relationships among those object properties. To manage their inquiries, the microworlds offer various tools to represent and manipulate objects. These representations allow properties and objects to be viewed from multiple mathematical perspectives.

We have developed several microworlds for mathematical inquiry. Each shares a common core of software. In all worlds, the system is used by students (usually in pairs) with a mentor present to coach the students and occasionally to provide seed topics or issues. The role of the mentor fades over time as students become comfortable with the software and acquire better inquiry skills. We have

done most of our testing in a lab setting with first- and second-year high school students, and junior-high students. In this paper, we will discuss the *Polygons* and *Graph Theory* microworlds only briefly (see McArthur and Lewis, 1991, and McArthur and Lewis, 1991b, for more information). We will begin by describing the structure of the worlds themselves -- the tools comprising the worlds and the objects students manipulate using these tools. Later we will give an overview of the kinds of knowledge students have acquired using our microworlds, including both their understanding of topics in polygons and graph theory, and of inquiry learning skills themselves. Finally, we will address the teaching and learning limitations of these worlds, and of ILEs in general.

The Polygons Microworld

In Figure 3 students are using *Polygons* to investigate the relationship of N and R (radius), as P (perimeter) is held constant. How is R changing (up or down) as N increases, and how fast? Students begin such investigations empirically, by generating several polygons (objects) and viewing them with different tools to find patterns. Objects

are created in two way. First students can input values for certain properties in an "object table", an instance of which can be seen in the upper-left corner of Figure 3. When two values are input the system computes all other values that logically follow, using equation solving and constraint propagation tools, borrowed from our the algebra tutor software. For example, if the student inputs an N of 8, then

polygons are created they appear on the large pictures window and can be manipulated in several ways -- they can be moved, queried and connected into different representations, as we describe below.

In the course of the inquiry shown in Figure 3, the students have used various tools for computing and representing information, including calculators, graphs and tables that

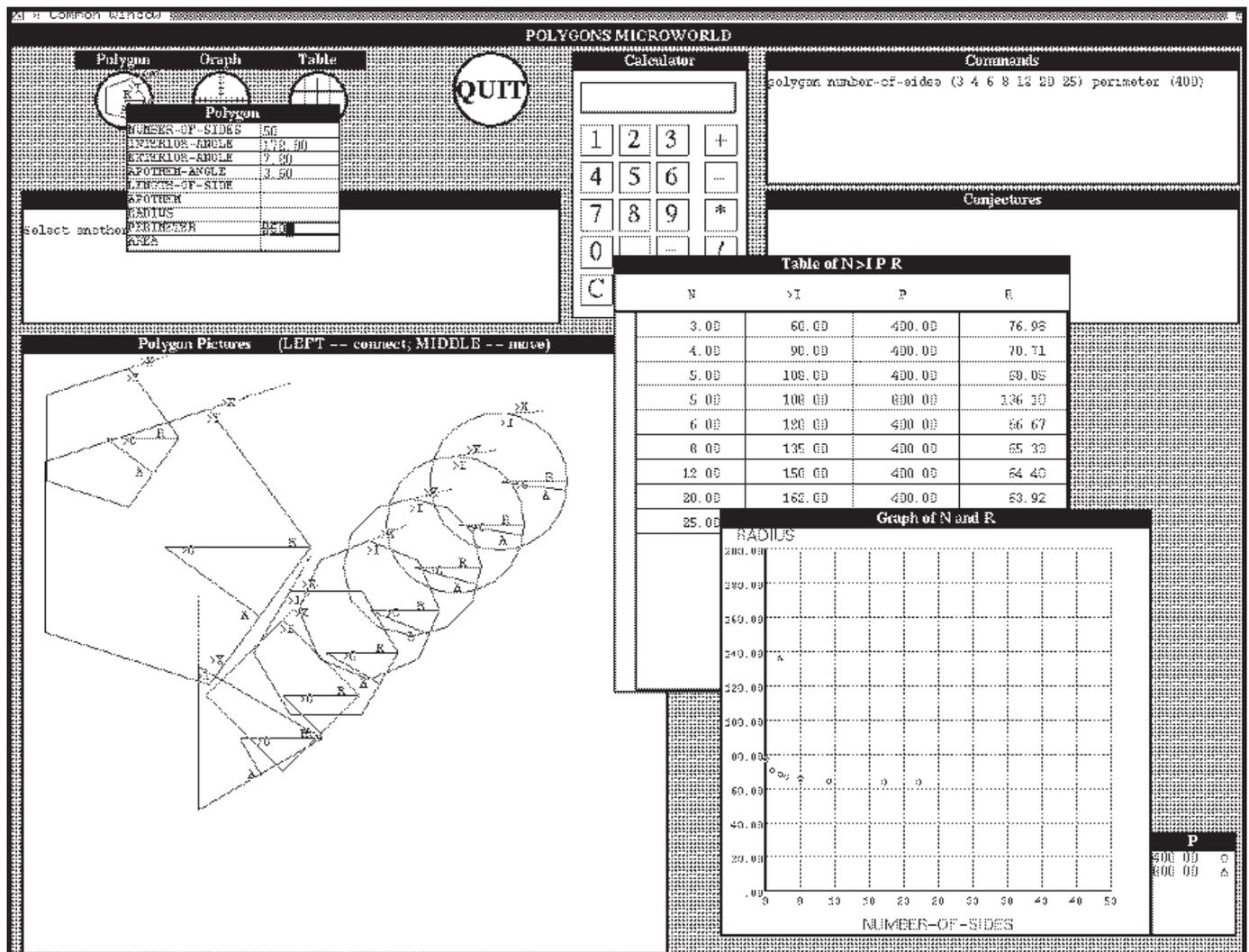


Figure 3: The Polygons microworld

>X (exterior angle; 45 degrees), >I (interior angle; 135 degrees), and >A (apothem angle; 22.5 degrees) are computed automatically. If N and some scale parameter (e.g., P) are both given, then all other values are determined and displayed. The second way to create polygons is to issue commands that will make many at a time, using the commands window in the upper right. When

show the relationship between N and R (in this case) from various perspectives. In Figure 3, the students have made a table that contains all the polygons, but represents only some of their properties. Tables and graphs are object-oriented -- each point or entry is an object, and a table or graph can be viewed as a filter or perspective on its objects. The object-oriented nature of the tables allows them

RESEARCH PAPERS

to be easily manipulated. Both tables and graphs can be moved and resized; tables can be sorted by any variable; and graphs can be "zoomed" to view a subset of objects.

By Figure 4 students have created an hypothesis about N and R and added it to the conjectures window (see the upper-right area of Figure 4). Our hope is that the investigation of one issue will trigger interest in new ones. In Figure 4 the student has suggested a new conjecture related to the one just completed: Since R decreases as N increases because the "points" of polygons are being "rounded off" (as the polygons approach a circle), the apothem (A) should increase to the same asymptote -- in other words, A should be equal to R as N increases to infinity. Testing this conjecture is straightforward, as shown in Figure 5. The student does not need to gather new data. Rather, the existing polygons simply need be viewed from a different perspective. The student creates this perspective

by generating a new graph for the properties N and A , and then "connecting" the graph of N vs. R to the graph of N vs. A . When one representation is connected to another, all the objects represented in the first are added to the second. In this case, then, the new graph immediately shows the relationship between N and A , and the student sees that her conjecture about N and A is confirmed.

Connection is a powerful tool. Each of our microworlds permit flexible connection among different representations. Any two data representations can be connected, including pictures; for example, you can point at a picture of a polygon and connect into it to a table. In addition, connection is incremental; after two representations are connected anything new appearing in one will appear in the other. Conversely, students can disconnect as well as connect representations.

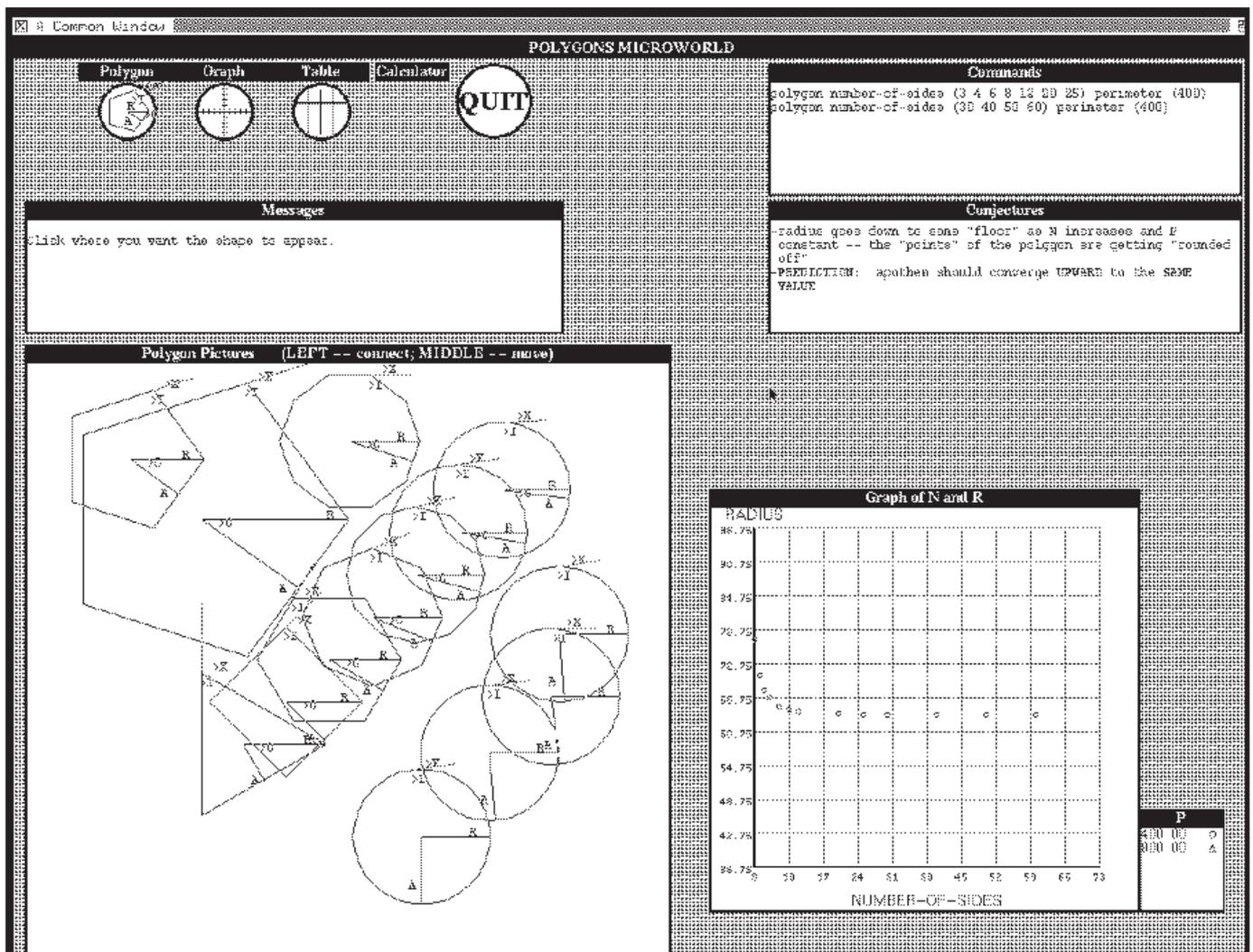


Figure 4: Connecting representations and making hypotheses in Polygons

The Graph Theory Microworld

Graph Theory is a more recent microworld through which students can explore the mathematics of linear graphs in discrete mathematics. This world is interesting for several reasons. *Graph theory* is an increasingly important area of discrete mathematics, and in recent years it has proven a valuable tool in gaming, modeling, and scheduling. NCTM standards (National Council of Teachers of Mathematics, 1989) advocate a place for graph theory and other areas of discrete mathematics in new curricula designs. However, few curricula available today include these topics, and the abstract character of many basic results in graph theory may make this a difficult topic for most high-school students. On the other hand, the visual nature of graphs suggests that a computer-based microworld might concretize these abstractions and empower students to explore graphs in productive ways. We are interested in

whether very young students acquire concepts in *graph theory*, and in seeing how microworlds support this learning. Thus, in *Graph Theory*, more than in *Polygons*, we are exploring new goals for student learning as well as implementing new methods of teaching and learning.

Analogous to *Polygons*, most inquiries in graph theory attempt to relate different graph properties, or to explain one set of properties in terms of others. In *Polygons* students work with properties like radius, apothem, and areas; in *Graph Theory* properties include elementary ones like vertices, edges, degree, degree sequence, and more aggregate properties like whether a graph is Eulerian, Hamiltonian, or planar. (NOTE: The degree of a vertex is the count of the number of edges incident to the edge; the degree sequence of a graph is a list of the degrees of all vertices in the graph; a graph is Eulerian only if all edges can be traced without retracing any; a graph is

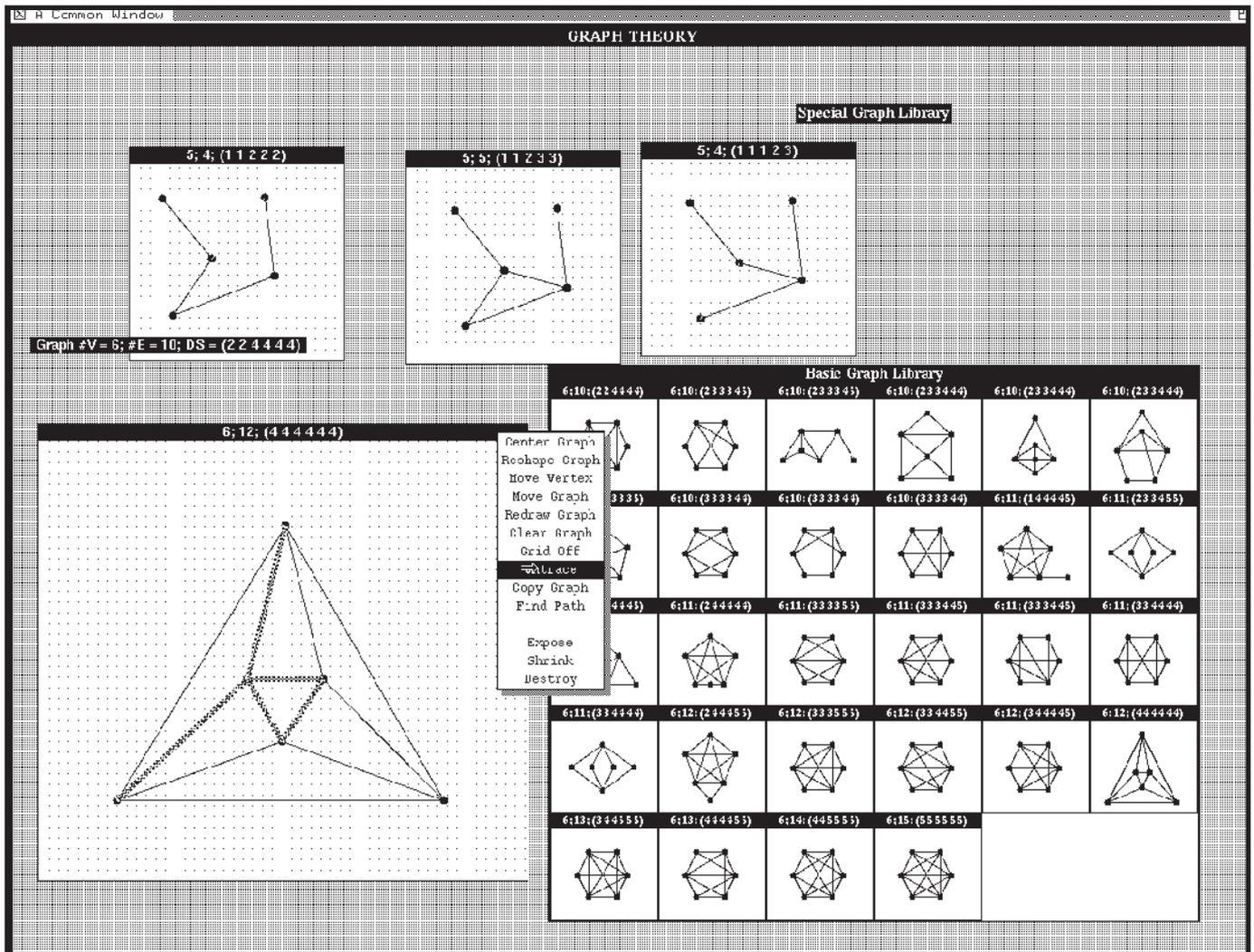


Figure 4: The Graph Theory microworld

Hamiltonian only if all vertices can be traced without retracing any; and a graph is planar only if the graph can be drawn using edges that do not cross.) As Figure 5 shows, students conducting graph theory inquiries use a variety of tools. They create and edit graphs using a mouse to add and delete both vertices and edges. Paths on graphs can also be traced and untraced interactively, or the student can request that the microworld find paths between any set of vertices. New graphs can be created by copying old ones and then editing the copy, rather than by making the new graph from scratch each time. Alternatively, graphs can be created by combining existing graphs or their complements. Graph libraries are also available for students to examine, and library graphs as well as student-generated ones can be used to suggest, confirm, or disconfirm student hypotheses.

The *Graph Theory* microworld shares many of the representations and features that characterize *Polygons*. Graphs, like *polygons*, are objects, and can be represented as pictures, as points in Cartesian graphs, or as rows in tables. The different representations offer different ways to view patterns of relationships among properties. For example, tables can reveal the relationship between number-of-sides and area in *Polygons*, and they are also useful to show how the number of edges and sum of degrees are related in *Graph Theory*. Connection between objects and representations in *Graph Theory* operates much as it does in *Polygons*. Pictures of graphs can be connected into tables or Cartesian graphs; tables and Cartesian graphs can be connected together; and representations can be disconnected. *Graph Theory* has one representation that has no analog in *Polygons*. Each graph has an associated list which includes the number of vertices, edges, and degree sequence of the graph. When graphs are created pictorially the system automatically updates these values, and changes are reflected on the black "title bar" at the top of each graph, as Figure 5 shows. As the Figure also shows, when graph pictures are "iconified" the icon identifies them by their vertex, edge, and degree sequence count. In effect, the graph picture and this list of key properties are distinct but tightly connected representations. The connection of these two representations, and the simple ability of *Graph Theory* to update the numerical properties as pictures are edited and modified, prove to be a simple but very powerful mechanism facilitating productive student inquiries.

Successes of Microworlds

In this section we briefly review the successes of ILEs in general, and of our microworlds in particular. Although many researchers are developing microworlds, there is still very little data to report. Most descriptions of microworlds discuss what an unfinished system will look like when it is completed. Of those that report a complete system, most describe how outcome data will be gathered in the future. And of those that present anecdotal data, most say how the system will be more rigorously tested at a later date. Only a handful of studies address the key questions concerning the inquiry-based methods of teaching underlying ILEs, and the goals for student learning towards which ILEs aim: Can students learn new and valuable concepts or topics through inquiry? Do students learn topic-specific skills effectively using microworlds or ILEs? Can they learn more effectively -- more rapidly or more deeply -- through inquiry than through instruction? What kinds of concepts or skills do inquiry microworlds encourage students to learn that cannot be easily acquired through ITS or through instruction or drill-and-practice methods of teaching and learning? And, do students inquiry skills themselves improve through the use of microworlds?

A series of studies by Shute, Glazer, and their colleagues represent the most ambitious attempt to date to answer some of these questions (Raghavan, Schultz, Glaser, and Schauble, 1989; Shute and Bonar, 1988; Shute, and Glaser, 1990; Shute, Glaser, and Raghavan, 1988). Of the three microworlds they have developed -- REFRACT (optics), VOLTAVILLE (electricity), and SMITHTOWN (macro-economics) -- the latter has been more extensively developed and tested. They found that students using SMITHTOWN acquired a better understanding of basic concepts in economics than did students receiving comparable classroom instruction (Shute and Glaser, 1990). However, they also point out that improvements were not uniform. In particular, some students demonstrated better strategies for inquiry and experiment planning, and their acquisition of economics concepts was, not surprisingly, also better. These differences in learning strategies and topic-specific acquisition were related to general measures of intelligence. Swanson (1990) notes a similar interaction of aptitudes and instructional methods. She found that inquiry-based

learning was less effective than direct lecture for less able students, while more capable students learned well using student-controlled inquiry. A conjecture worth further testing is that learning through inquiry is both rewarding and demanding. Bright students may be able to acquire knowledge more deeply and rapidly, but to do so they will require scientific inquiry skills that themselves need training.

Harel and Papert (1990) have reported comparable successes in the instructional software design project (ISDP). Students learned about fractions using computers and LOGO in the context of a relatively rich and unstructured learning environment that might be loosely regarded as a microworld. Harel and Papert claim that students working in ISDP acquired more knowledge about fractions than students who were taught in a lecture format. Unfortunately, their study cannot be used to directly compare inquiry to instruction methods of learning, since students using ISDP had an unfair advantage over control students -- they learned using ISDP and received instruction while the control students received only instruction.

However, this limitation probably does not greatly concern Harel and Papert, since they do not appear to be very interested in comparing how different methods of teaching promote learning of traditional subject-specific knowledge. In addition to fostering a procedural understanding of fractions, they have two more novel outcomes in mind: (i) helping students acquire a deeper understanding of the concept of fractions and how fractions relate to objects in the real-world; and (ii) supporting the development of important metacognitive (inquiry) skills.

To measure learning of traditional fraction skills Harel and Papert naturally borrowed existing diagnostic tests. But to assess how well ISDP accomplished its more novel learning goals, they used "thick descriptions" of students interactions with ISDP, since no standardized instruments can currently measure these kinds of outcomes. Thick descriptions are rich qualitative analyses of students behaviors, plans, and intentions (Geertz, 1973). Using these techniques, Harel and Papert found, among other things, that by moving between multiple representations in ISDP students acquired deep knowledge of rational numbers, that students in ISDP learned LOGO better than comparison students, and that ISDP students developed good metacognitive skills.

Like Harel and Papert, our initial work has attempted more to characterize the outcomes of learning through inquiry rather than to compare inquiry and instruction. We agree that outcomes like deeper conceptual understanding and improved metacognitive skills are important, and we are trying to operationalize these ill-defined ideas by describing specific inquiry learning processes and products at a detailed information processing level. Here we will only briefly summarize our observations; more extensive discussions can be found in other papers (e.g., Lewis, McArthur, Bishay and Chou, 1992; McArthur and Lewis, 1991; McArthur and Lewis, 1991b; Lewis, Bishay and McArthur, 1993, 1993b).

Learning expected concepts.

Using both *Polygons* and *Graph Theory* students were able to generate and resolve (often with substantial mentoring) a wide range of inquiries. In many cases inquiries began as unspecific and qualitative. For example, in Figure 3 and 4 students first focused on conjectures that considered how R changes (up or down?) as N increases. In later sessions these initial results often laid the foundation for more specific quantitative inquiries, such as developing a function that relates R to N ($R = 180 - 360/N$). Over several sessions, then, students learned many of the standard results and theorems that can be found in texts on the geometry of *polygons* or *graph theory*.

Learning unexpected concepts.

Since students were usually given relatively free reign in directing inquiries, we hoped that some issues and hypotheses would arise which we had not foreseen. We did indeed find that inquiries often took unanticipated turns. For example, one student developed a series of increasingly general hypotheses about Eulerian graphs (What characteristics of a graph make it possible to begin at any vertex and trace all edges of the graph, returning to the original vertex without tracing any edge more than once?). At one point the student was attempting to verify that graphs with just one vertex of odd degree could not be so traced. (NOTE: The degree of a vertex is the count of edges coming into the vertex; hence a vertex of odd degree has an odd number of incident edges.) In doing this she noticed that none of the graphs she had created, nor any in the graph libraries, had just one vertex of odd degree. This datum was sufficiently interesting to cause her to suspend the investigation of Eulerian cycles and define a

new issue to investigate the lack of graphs with just one vertex of odd degree. This inquiry ended with the student discovering and explaining the *handshaking lemma* for simple graphs. (NOTE: The handshaking lemma states that in any graph the sum of all vertex degrees must be equal to twice the number of edges, hence must be even.) While this issue was both interesting and important, it was certainly not one that the mentor had anticipated ever broaching.

Some inquiries were even more surprising. In one session with *Polygons*, for example, the student quickly established an equation relating N and the apothem angle ($>A$): $>A = 180/N$. To establish this rule the student successfully used the strategy of multiplying ($>A * N$ always yields a constant value). In his next inquiry, he tried to apply this strategy, or a similar one, to describe the relationship between N and $>I$. Attempts to generate a constant by multiplication ($N * >I$), or division ($N / >I$) were not successful. However, when the student tried $>I / >A$ he found it yielded $N - 2$, leading to the equation $>I / >A + 2 = N$. This equation was not part of the mentor's informal agenda of topics; indeed, it came as a complete novelty to the mentor as well as the student. The discovery is one of the infinite number of theorems that never find their way into textbooks, but which any student can discover, name, and own -- in our sessions, for example, the above result became known as "Jody's Theorem". Such discoveries allowed students to understand that they can create their own mathematics, not just learn what others have found.

Opportunistic learning of concepts outside the microworlds.

The above examples underscore the *opportunistic* nature of inquiries in *Polygons* and *Graph Theory*. Both worlds proved sufficiently dense in interesting issues that, in addition to unanticipated ideas within geometry and *graph theory*, some concepts complete outside both areas occasionally took center stage. For example, in *Polygons* students investigated the concept of limits, different types of variables or scales, the distinction between independent and dependent variables, the notion of a mathematical basis, and constraints that bound the application of equations when modeling real-world situations. We generally encouraged the students to pursue these issues if they seem inclined. In contrast, the microworlds developed by Shute and Glaser maintain relatively well-defined and pre-defined objectives for student learning.

Learning generic and metacognitive concepts.

Working with the microworlds also permitted students to acquire some rather broad insights into the nature of mathematical knowledge and problem-solving. As we just mentioned, students began to understand that they could make their own mathematics, rather than learn about theorems discovered by others. They also learned several other general lessons. For example, as students built quantitative hypotheses on previously confirmed qualitative ones they slowly abandoned the notion that mathematics comprises a series of separate problems to solve. Thereafter, they often revisited previous hypotheses -- sometimes several sessions later -- and incrementally refined them or built new ones from them. In short, the students were able to use the microworlds to learn various topic-specific (geometry or *graph theory*) ideas. But as important as each individual idea was the broader understanding of how these ideas interconnect and build on one another.

Comparing ITS and ILEs

Depending on your perspective, today's ILEs and microworlds may look more successful than ITS, less successful than them, or their successes may be incomparable. How you judge them depends on the outcomes you value in student learning and what you count as evidence that important learning goals have been reached. ILEs enjoy modest success insofar as they offer at least anecdotal evidence supporting some of the broad claims for inquiry-based learning methods in the literature (see Bruner (1961), and Ausubel (1961) for discussions of some of the purported advantages of learning through discovery in theory and practice). ILEs that empower students to create their own mathematics can help them acquire traditional subject-specific knowledge; for example, students can discover the theorems like the handshaking lemma that we see in any text on *graph theory*. There is also some evidence that well-crafted microworlds can make this subject-specific knowledge accessible to younger students -- *graph theory* is rarely learned by middle-school students, for example -- and perhaps can improve the rate at which concepts are learned, at least by above-average students. Microworlds and ILEs may help students acquire a deeper understanding of concepts; or at least an understanding that is different from that typically communicated in texts. And rich exploratory environments may expose students to

unanticipated ancillary concepts, like limits. In addition, students using ILEs may improve some inquiry skills or related metacognitive skills which they rarely have a chance to practice in more teacher-controlled curricula. Finally results like "Jody's Theorem" indicate ILEs can positively influence affect, motivation, and the perceived "ownership" students feel towards mathematical ideas and practice.

However, to our knowledge no microworld has demonstrated "2 sigma" improvements in student outcomes on standardized tests that some ITS can boast. In general, ILEs have not demonstrated yet they can yield the improvements in student learning that are associated with one-on-one tutoring. But, other than Shute and Glazer's work (1990), we see little interest in such "horse race" evaluations. Instead, there has been a realization that many important learning outcomes are not central goals of educational software, and most ILEs thus attempt to articulate these new goals and design software to facilitate learning implied by these goals. Since they differ on goals for student learning, then, direct comparison of ILEs and ITS is difficult if not impossible.

For similar reasons, comparison of learning outcomes with ILEs to outcomes without ILEs are as difficult as comparison of ILEs and ITS. "Before-and-after" evaluations of ILEs require control classrooms that can provide baseline data on how effectively the learning goals ILEs target are usually accomplished. Ideally, control and experimental classes would differ just one way -- keeping the learning goals of the classes and the methods of teaching constant, the experimental class would use the ILE and the control class would not. Only under these tightly controlled circumstances could you confidently conclude that ILEs caused improvements in learning outcomes. But such control classrooms may not exist, since most ILEs do not try to improve the effectiveness with which established educational goals are accomplished, nor do they advocate traditional methods of teaching. In later sections we shall address these evaluation problems in more detail.

Reasons for ILE and microworld successes

It is important to account for the successes microworlds and ILEs enjoy (or the successes they may have in the future) for the same reasons ITS should be subjected to such analysis. Each microworld is no more enduring than an ITS. Its enduring contributions lie in the general principles

it suggests, confirms, and disconfirms. However, several challenges face any assessment of the reasons for ILEs' real and potential successes. As just discussed, there is little data from which to generalize, and the data that is available is ambiguous at best. Clear consensus has yet to emerge about what students are learning or should be learning, so it may be premature to speculate in any detail on why ILEs and microworlds enhance that learning. With these qualifications in mind, in this section we attempt a brief and somewhat speculative account of some reasons ILEs may enhance learning, in the future if not today.

Many plausible reasons that ILEs may contribute to learning are supported as much by theoretical arguments as data. For example, we earlier suggested that ILEs are often justified by a constructionist (versus instructionist) view of learning, and by a generally high regard for the value of student controlled learning and feedback. These arguments have been elaborated elsewhere (see e.g., Bruner 1961; Harel and Papert, 1990; Papert and Harel, 1991). Below we take a slightly different perspective, focusing on a few reasons for success that our microworlds underscore.

Knowledge used for delegation and role sharing, not tutoring.

Like ITS, ILEs often embed considerable intelligence, although it is less obvious, distributed in different ways, and used for different purposes. For example, in *Polygons* and *Graph Theory*, knowledge is embedded in several tools, including the constraint equations (*Polygons*) and collections of predicates, functions, and deductive inferencing capability (*Graph Theory*). Essentially, these are parts of an expert system for these subjects which have been broken out into distinct components. In some cases the tools are not used by the student directly, but provide the powerful automatic capabilities. For example, the constraint propagation functions of *Polygons* allow students to create any possible regular convex polygon. This gives students greater flexibility and control of object creation. It also allows students to investigate a much wider space of issues than pencil and paper would permit. In *Graph Theory* various predicates can be invoked as tests by the student. For example, students can just ask if a graph is Eulerian; they do not have to apply this time-consuming and error-prone procedure by hand. This is useful if main goal of learning is to discover and understand patterns in mathematical objects rather than to rehearse procedures.

In general these intelligent embedded tools permit some less important computation or supporting cognition to be *delegated* the software. ILEs and microworlds thus implicitly or explicitly divide the cognitive skills in an area of inquiry into various packages. Some packages, like procedural expertise involved in solving equations and testing graphs for Eulerian cycles, are regarded as less important for students to learn -- or students are simply assumed to have mastered them. However these skills often need to be invoked when conducting any inquiry that practices "higher level" skills that today are viewed as more valuable to learn. For example, although tracing Eulerian cycles is a routine procedural skill, to reason at a "high level" about what makes a graph Eulerian you need to be able to generate such cycles. This creates an apparent paradox. In any rich inquiry, how can you arrange for students to focus on the important skills when relatively unimportant ones must also be applied and may require much of the cognitive effort? This is a long-standing dilemma in pedagogy.

In traditional mathematics texts and curricula, this dilemma is often solved by carefully crafting a series of problems that are stripped of all but one important concept. For example, we often see a collection of problems to practice the distributive law that differ only in the numbers they use; then a similar set to practice substitution, and so on.

ILEs and microworlds offer a different answer. Instead of requiring that students negotiate relatively stripped and carefully ordered problems to ensure that the students practice key skills and minimize time spent on less important ones, ILEs generally permit the student to choose any problem or inquiry he or she wishes, however complex. The student's practice is focused not by crafting a series of problems but by automating many of the less important or clerical skills, delegating them to the software. In this way, the student is free to focus on planning inquiries, or testing conjectures, rather than spending most of his or her time filling in tables or drawing polygons and graphs. ILEs and microworlds thus focus students' learning implicitly by what they choose to delegate and what they leave up to the student, rather than explicitly by generating carefully sequenced problems.

Delegation as a technique to enhance student learning has several potential benefits, although few have been explored systematically in research. Like carefully

developed sequences of problems, delegation implements the principle of providing *intensive and focused practice* on to-be-learned skills. However, unlike stripped problem sequences, delegation can provide focused practice that is also situated or *contextualized*. In mathematics texts, problems are stripped to a few concepts to facilitate intense practice on one of two ideas. But in stripping down problems, they are typically robbed of related concepts that are often connected with them in more realistic or authentic problem contexts. Through delegation students are able to work immediately on problems or issues that embed several kinds of concepts or tasks. They may practice only one or two concepts at a time, but these concepts are always situated in a context with other tasks; ones that are delegated to the microworld or ILE itself.

Extending students' learning capabilities through cognitive amplifiers.

Viewed from a slightly different perspective, delegation has an additional potential value. In microworlds and other ILEs the system is expert in relatively rudimentary skills and can be delegated tasks that require them. Because the students are thus freed of these elementary tasks -- the ones that typically ITS attempt to tutor -- ILEs effectively permit students to investigate new, possibly more highly valued, topics in mathematics. For example, in *Polygons* students may create and confirm their own theorems rather than apply published formulas to compute area or perimeter; and using *Graph Theory* they can actually understand some basic results in discrete mathematics rather than work through specific traveling salesman puzzles. In general, microworlds can be viewed as inquiry partners, and the division of labor permits the machine do utility procedures while enabling students to learn more complex ideas that would otherwise be conceptually beyond them. In this sense, ILEs can be viewed as cognitive amplifiers or magnifiers.

Resnick's (1991) microworlds built on *LOGO provide an excellent illustration of this principle. *LOGO is a distributed version of the LOGO programming language that permits learners to write rules governing the behavior of thousands of interacting computer "creatures" -- for example ants searching for food or cars in a traffic jam. Students then observe the group behaviors that emerge when these populations are run in a simulated world, exploiting parallel supercomputers that associate each creature with a

distinct processor. In this partnership, then, *LOGO knows everything about the creatures at a local level, however the microworld knows nothing of the global patterns of group behaviors that emerge. The microworld generates these behaviors, but it is the student's goal to characterize the emergent features and to try to relate them to local behavioral rules. On the other hand, while students are asked to characterize emergent behaviors, generating the behaviors is beyond any unaided human. Thus delegating simulations to *LOGO creates more than just a partnership that allows students to learn new ideas that were previously difficult for them. The partnership permits learners to gain insights into phenomena that are impossible to understand without such tools. Indeed, similar computer-based environments are also essential tools for mathematicians and scientists who seek to understand non-linear and chaotic systems (Gleik, 1987). Chaos as a field of study could not exist without computers. Here, then, is an example of how microworlds can open up goals for student learning that are not merely new to high-school curricula, but new to curricula at any level.

Decoupling ILE expertise from student learning.

The knowledge embedded in ILEs thus contrasts with in an interesting way with the expertise in ITS. While the expert systems in ITS represent the knowledge that students will learn, many ILEs take as much care in representing the skills that will be delegated to the system and that will therefore not be the main focus of student learning. Of course, ILEs and microworlds must also be crafted to elicit and support the knowledge students will learn. For example, in *Polygons* students focus on developing hypotheses and refining them, and the software provides tools through which hypotheses can be expressed. Nevertheless, most of the programming complexity in our microworlds concerns tasks delegated to the software, not tasks and knowledge delegated to the student.

In general ILEs and microworlds are freed of a heavy demand that burdens ITS -- to represent much, if not all, of the knowledge of the "ideal" student. It is critical for ILEs and microworlds to represent much supporting knowledge, and to supply tools that students can use to construct their own knowledge. But the central premise our ILEs is that good learning environments need not be expert in the skills they are trying to help students learn. Since ILEs embed less knowledge than ITS, it has been easier to develop relatively successful ILEs in subjects and topics for which ITS cannot

yet be built. Further, as we just noted, some of the learning outcomes associated with ILEs have not even been effectively operationalized. If we cannot assess these skills we certainly are not in a position to develop ITS that can model and coach them.

This generalization does not apply to all ILEs. Several systems described as microworlds do embed extensive knowledge of the subject which students are to learn, and some knowledge of discovery or inquiry itself. For example, SMITHTOWN (Shute and Glaser, 1990) represents both economic concepts and rules for scientific inquiry, and attempts to remediate misconceptions in students' inquiry skills. Similarly, White and Frederiksen (1986) describe an ILE for electronics which relies on a carefully crafted progression of qualitative models, meant to mirror successive stages in students' understanding of circuits. However, unlike other ILEs we have reviewed, including our own, these systems have relatively well defined goals for student learning and may manage learning more like an ITS than like an inquiry-oriented ILE. Indeed, such systems represent an interesting mixture of ILEs and ITS, about which we shall have more to say later.

Limitations of Microworlds

In spite of their promise and current popularity, computer-based microworlds and ILEs enjoy few large-scale successes to date. Microworlds that support student-centered learning through inquiry are certainly consistent with many calls for curriculum reform. However, these ideas and systems are not in classrooms on a large scale. Here we will discuss three distinct challenges facing ILEs. We begin with cognitive challenges -- the inherent complexities of learning through inquiry that future ILEs should attempt to address. But two other limitations cannot be solved by purely technical means. *Evaluation* challenges facing ILEs concern the learning outcomes and goals they aspire to. As we have noted, it is very difficult to operationalize many of these goals and to measure their learning outcomes using standardized tests. And *implementation* limitations involve the multiple complexities of realizing an inquiry method of learning in classrooms on a broad scale.

Cognitive problems: Large search spaces and thrashing

One of the main complaints voiced against learning through inquiry is that it is an inefficient way to acquire knowledge. This criticism is certainly not new with ILEs and

microworlds. For example, Ausubel (1961), among others, criticized unguided discovery learning by pointing out how much of subjects' behavior appeared irrelevant to the learning goals of such studies. He claimed that there are faster ways of learning specific skills, at least when the skills can be clearly defined. A more modern description of this inefficiency is that relatively unguided inquiry challenges students by placing them in a large search space (Newell and Simon, 1972; McArthur and Lewis, 1991, 1991b) in which only a relatively few choices represent useful directions for investigation.

Microworlds and ILEs do address the problems of inefficiency in some respects. Many of the tools that comprise such environments can be viewed as ways of speeding up the processes of inquiry. For example, when students using *Graph Theory* delegate the testing of Eulerian cycles to the system, or expect the system to find paths meeting certain conditions, they both accelerate inquiry and reduce the chances of unprofitable procedural errors. In effect, delegation provides students with cognitive amplifiers that will help students overcome many inefficiencies in searching large spaces of issues, problems, hypotheses, and data.

Nevertheless, ILEs also potentially add to the problems of large search spaces and inefficient learning at the same time as they provide tools to help ameliorate them. If students can search spaces faster using ILEs, they also have bigger spaces to search. For example, in *Polygons* by giving a value for any single scale and angle variable -- for example, N (number of sides) and P (perimeter) -- students can create virtually any regular *polygon*. Working in a pencil-and-paper microworld, they would be much more constrained in the data they could generate. Thus students using *Polygons* can generate data more rapidly than student using pencil and paper. But they also have to plan their data gathering more carefully, since they are more ways here to gather data inappropriate to an issue or hypothesis.

We have begun to analyze several kinds of search inefficiencies our students demonstrated in *Polygons* and *Graph Theory*. Some problems relate to weaknesses in students' *experiment planning skills*. For example, when investigating what kind of *polygon* maximizes area for a given perimeter, one student began generating *polygons* with area held constant and N varying. Yet this type of mistake was relatively rare. More common were problems

in managing different tasks or issues. For instance, a pair of students using *Graph Theory* might begin by attempting to investigate the properties common to all trees, then, without completing this issue, find an interesting pattern in a graph that had cycles. (NOTE: A tree is a connected acyclic graph -- a graph with no cycles where all vertices are connected, directly or indirectly.) This would then be the seed of a new issue and the first issue would never be resumed. In turn, the second issue might suffer the same fate. We refer to this pattern of search behavior as "thrashing". Elsewhere we describe in detail the topography of "thrashing", "data-dominated", and "hackerly" reasoning and other search problems (Lewis, McArthur, Bishay and Chou, 1992; McArthur and Lewis, 1991; McArthur and Lewis, 1991b).

We expect this analysis search skills and problems will have several benefits. First, it has been argued that many apparent inefficiencies in inquiry learning mask subtle but valuable learning benefits (Papert 1980, 1993). Part of this debate may be a definitional problem -- as we have noted, what constitute desirable outcomes of learning through inquiry remain unclear. However, we believe that various inefficiencies associated with inquiry learning can be operationalized, and we can thus look for some of the potentially hidden dividends of apparently inefficient inquiry. Second, this detailed analysis will provide a basis for designing new microworld tools and supports that can help students overcome often very predictable problems in their inquiry and search skills, while not taking initiative away from the learners. We regard this as probably the central issue for improving ILEs in the future, and as a key technical challenge to making inquiry a viable method of learning.

Evaluation problems: Determining and measuring learning outcomes

One of the major limitations of ILEs that we noted briefly has little to do with software difficulties or with problems in student learning through inquiry. It is the problem of defining the goals for learning or learning outcomes that the ILEs aspire to, operationalizing these outcomes in objective instruments, and applying these instruments to assess the efficacy of the microworlds in promoting their learning goals. A few ILEs, notably SMITHTOWN, are attempting to improve students' performance on well-defined and clearly operationalized skills and knowledge. However, as we have noted, other ILEs have multiple goals for learning, and less clearly defined ones. In general, in

spite of a broad interest in inquiry-based learning, today we have no clear consensus on what kinds of skills we want ILEs to engender, let alone objective standardized instruments to measure these skills.

In some cases, (e.g., learning *graph theory*) we lack effective instruments simply because these skills are not part of existing curricula. Presumably, it will be straightforward to develop tests that measure this knowledge. But other skills will pose substantial challenges. Some knowledge (e.g., learning creative inquiry skills or understanding that new mathematics can be discovered, not just taught) may be inherently "fuzzy" and very difficult to measure with traditional multiple-choice or short-answer tests. Other important new outcomes may be well-defined products -- for example, the novel theorems our students discovered using *Polygons*. Yet, since this learning is opportunistic, these outcomes will also not be uncovered by standardized tests which typically assess a fixed set of learning goals. In general, quantifying and evaluating such outcomes is very much at the formative stage.

Thus, today ILE developers face several interrelated challenges. They are creating computer-based environments that implement an inquiry-based method of teaching and learning, at the same time as they are trying to demonstrate that such methods, effectively implemented, can lead to highly desirable student learning outcomes. However, only in a few cases are these desirable outcomes well-defined and comparable to outcomes towards which ITS aspire and which are topics in most curricula today. Therefore, developers of ILEs often need to operationalize important learning outcomes and to devise new instruments for assessing the successes or failures of their systems. Since they cannot borrow standardized tests, they essentially must play the dual roles of curriculum developer and evaluation expert, as well as system builder.

Even if we can develop relatively valid and reliable measures of important skills learned through inquiry, it may be difficult to pin down in any detail the role of ILEs in learning. As mentioned, ITS can often be placed in classrooms with little disruption of ongoing practice, facilitating relatively controlled comparisons of outcomes with and without the ITS. (NOTE: However, see Schofield, Evans-Rhodes and Huber (1990) for one in-depth analysis of an ITS which reveals that the system triggered several enduring changes in classroom practice. These changes,

as much as the ITS itself, may be responsible for outcome improvements. Causal attributions are often complex even in the simplest and most controlled tests of educational software. Nevertheless, the changes reported represent relatively small modifications that leave the basic structure of the classroom, and roles of teachers and students, relatively intact. ILEs, if successful, may demand more fundamental changes.) Moreover, some ITS permit an even finer grained analysis the effects of individual software modules within the system. In "ablation" studies one component of an ITS (e.g., student modeling, some interface feature, or various pedagogical rules) can be disabled and student learning with the ablated ITS compared to learning with the complete system. Using ablation students, ITS developers can, for example, determine what role student modeling plays in learning. It may even be possible to use such results to speculate about such modeling in human tutoring -- since human tutor ablation studies are probably impossible!

On the other hand, when ILEs are tested in classrooms or in labs, the software itself is never the only change. Because most ILEs attempt to change learning outcomes and to implement methods of teaching and learning that are novel to classrooms, they often demand fundamental classroom restructuring. As Harel and Papert (1990) have noted, microworlds are usually embedded in larger framework of changes or reform that can include new curricula, novel means of evaluation (e.g., projects versus tests), different teachers roles, and an altered rhythm of work. Thus, ILEs are rarely subjected to experiments that would systematically alter one variable at a time to unveil the causes of improvements in learning outcomes. As a result, it is difficult to make a strong case for the particular roles that ILEs alone play in improving student learning outcomes and to defend the importance of inquiry-based methods of teaching and learning they embody.

Implementation problems: Demands on the classroom and teachers

The fact that the new methods of teaching and learning associated with ILEs usually usher a broad set of changes into the structure of classrooms and curricula implies implementation problems as well as evaluation problems. Indeed we suspect the practical problems of implementation may dominate cognitive and evaluation difficulties in determining whether ILEs in particular, and inquiry-based learning in general, will enjoy any wide-scale

success in classrooms. Some research has already begun to document how microworlds and ILEs will change the classroom culture. Harel and Papert (1990) describe extensive changes to the physical layout of classrooms, the culture of learning, and the transition of teachers from instructors to guides. And Schofield, Evans-Rhodes and Huber (1990) documented how Anderson's Geometry tutor caused shifts in teachers' attention to different types of students, in teachers' classroom roles, as well as in the effort and involvement of students.

Our own experience is consistent with these results. Traditional ITS, like our algebra tutor, have required little implementation effort, and almost no new curriculum development. Thus, roughly 75% of our time in fielding ITS has been spent developing the software, and only 25% of our time involved classroom implementation. On the other hand, in fielding *Polygons*, *Graph Theory*, the time required to develop software and implement it in the classroom was roughly reversed. Software development took 25% of our time, while implementation and curriculum development required about 75% of our effort. Our student-centered courses in mathematical modeling (Robyn, Stasz, McArthur, Ormseth and Lewis, 1992) and statistics (McArthur, Robyn, Lewis and Bishay, 1992) also reflect this division of labor. In both projects, the cost of implementation dominated the cost of technology development (Berman and McLaughlin, 1978).

In the future, this allocation of labor is likely to be the rule, not the exception. Our courses are not exceptionally novel. In fact, they are consistent with new NCTM curriculum standards (National Council of Teachers of Mathematics, 1989), in content and organization. To the extent that these emerging standards take hold, many courses will mirror the structure and content of our courses in a few years. NCTM's and professional standards (National Council of Teachers of Mathematics, 1991) do anticipate some of the difficulties that teachers will face when trying to organize curricula around inquiry and discovery. Nevertheless, understanding the roles teachers will need to play, and getting these roles into widespread practice, remains a big challenge.

The challenge becomes more acute and important when we realize how much at odds new methods of inquiry-based learning are with current and past classroom practice. While ILEs have recently become increasingly

popular in research and in the lab, the principles upon which they are based are not new. As Lawrence (1970) notes, perhaps as far back as Plato scholars have hinted at the value of discovery in improving retention (Spencer and Seneca), deepening understanding (Rousseau and Kant), and enhancing motivation (Wyse). Early this century, Dewey's progressive education movement resurrected and refined these views, advocating active, constructive, and student-centered learning. The progressive movement flourished briefly, but by the end of the depression its impact in schools was clearly marginal. The next cycle of interest in inquiry-based learning was largely triggered by Sputnik, and the perceived threat of Soviet scientific supremacy. Bruner (1961), among others, eloquently championed the importance of learning through discovery, and offered some empirical evidence of its efficacy. However, the new movement had equally eloquent critics (e.g., Ausubel, 1961), and by the end of Learning by Discovery conference in 1966, it had already begun to unravel in empirical and definitional disagreements.

From an historical perspective, then, interest in inquiry-based learning has been cyclic. The recent calls for an increased emphasis on discovery and inquiry in learning represent at least the third time this century these ideas have moved into the educational research limelight. Among others, Cuban (1986) and Cohen (1988) have chronicled the history of these cycles. Cuban notes that, while the popularity of these ideas in the classroom has waxed and waned, the trend-line still remains firmly away from inquiry-based learning and towards instruction-based pedagogy. The obvious question is whether this cycle of interest will again subside with little lasting effect on classroom practice. Technologies like ILEs represent the only new factor in the equation; the hope is that they can catalyze an enduring change in the trend-line of pedagogy. But Cohen cautions that the dominant view in education, and in our culture in general, continues to be "teaching is telling and learning is listening" (Cohen, 1988). Unless that fundamental view changes, ILEs and microworlds will remain a marginal tool in learning, relegated to lab studies and a few well-publicized demonstration schools. On the other hand, ITS, which may not look as attractive as ILEs in theory or in the lab, may fit much better with existing classroom practice and with prevailing attitudes about learning and teaching.

Conclusions and speculations about future roles of AI in education

This section will summarize and draw together various ideas scattered through preceding parts of the paper. In doing so we will also extrapolate from current trends, and from the noted strengths and weaknesses of ITS and ILEs, to project some possible future roles for artificial intelligence and knowledge-based systems in education.

Increasing diversity of applications of AI in education

ITS and ILEs are often discussed as if they were the only two distinct approaches to developing knowledge-based systems for education. On the surface, this paper -- discussing first ITS then ILEs -- perpetuates that image. However, on a more careful reading, it should be apparent that a field which began with two relatively clearly opposing positions has now begun to fracture into a multitude of related systems and approaches. Instead of two points of view on the application of artificial intelligence to education, it is more accurate today to see the field as a *continuum*, with omniscient, tutor-controlled ITS on one end, and completely student-controlled ILEs on the other.

The *continuum* represents a variety of different ways of dealing with the problems and weaknesses of the two opposing approaches, outlined in the previous sections. In general, the weaknesses of one approach can often be creatively confronted by borrowing some of the strengths from the opposite extreme, in effect creating a range of systems. Such systems may blend various kinds of tutor- and student-controlled activities. Consequently they are beginning to resemble one another in structure and design, although they often claim distinct conceptual lineages -- for example "constructionist" versus "instructionist" theoretical foundations -- associated with the extreme end points (Papert and Harel, 1991).

At the ITS end of the continuum, one major problem is that a thoroughgoing intelligent tutor must be highly knowledgeable about the subject it tutors. For every problem it poses to the student, or every request for information and help from the student, the system is expected to know the "right" answer. This demand for omniscience limits ITS to tutoring subjects for which we have relatively complete cognitive task analyses, including not only an understanding of the competence of the "ideal"

student, but also the misunderstandings of the novice. However, there are few interesting subjects, from an educational perspective, for which such complete task analyses are available. Indeed, as we have suggested, subjects for which such analyses are possible may become uninteresting to teach simply because they have been fully analyzed and hence can be completely delegated to machines.

In response to this dilemma, many researchers have begun to consider ways of weakening the assumption of omniscience, to permit the development of at least partially intelligent tutors for subjects or skills that are regarded as more valuable. For instance, although we classified SMITHTOWN (Shute and Glaser, 1990) as an ILE, it actually combines some features of ITS and ILEs. It embeds a rudimentary expert system for scientific inquiry skills, represents some of common misconceptions about planning scientific experiments, and it can model students' scientific skills at least to the point of recognizing examples of good planning and common planning "bugs". For example, if a student insisted on changing the value of several variables at once, instead of manipulating one dependent variable at a time, SMITHTOWN will coach this tactical error. Clearly SMITHTOWN does not pretend to embed a complete cognitive analysis of scientific inquiry skills. Rather, the expert system is relatively skeletal, and the catalogue of student misconceptions is equally incomplete. Nevertheless, this "semi intelligent" tutoring system can be very useful in helping students manage their inquiries. In this sense, SMITHTOWN represents a way of moving the principles of ITS into new topics of learning -- here, learning about inquiry itself -- that have high educational value.

At the ILE end of the *continuum*, a central technical problem is the challenge of large search spaces. We have noted various apparent inefficiencies associated with student-controlled navigations through large collections of issues. While acknowledging that some of this thrashing and floundering may yield hitherto undocuments benefits, it is also important to consider techniques for gently curtailing unprofitable inquiries and guiding them in more useful directions. Implicitly or explicitly, most ILEs provide guidance in one form or another. For example, in *Polygons* and *Graph Theory*, menus of object properties -- area, number-of-sides and perimeter, for example -- will encourage students to examine relationships between

these variables, and will discourage investigation of other variables or ideas not mentioned at all. By reducing the set of variables explicitly mentioned in the menus, the microworlds can easily control the "dimensionality" of the students' search space (see McArthur and Lewis, 1991, 1991b for a detailed discussion of other passive techniques to help manage search, and of the more aggressive approaches briefly noted below).

However, some ILEs are now considering more active ways of guiding students' inquiries. We have used our analysis of students' problems with search and thrashing in *Polygons* and *Graph Theory* to design a collection of heuristic rules that will intervene and make suggestions about how to redirect an inquiry, if the microworlds detect patterns of reasoning that appear particularly unprofitable. For example, (as in SMITHTOWN) if the student insists on changing several variables at once, the microworld will question this behavior, although it will not prevent such decisions. Or, if the student switches from one inquiry to another without completing the first, the microworld will mention the transition, although it will not demand that the student return and finish the first issue.

Mixed-initiative systems and locally intelligent agents

There are several points to make about these examples. First, the rules by no means represent a global cognitive model of inquiry in our microworlds. Rather, they are *locally intelligent agents* that impose islands of tutor-control in a relatively large expanse of student-controlled inquiry activities. The local interventions are carefully chosen so that student initiative is not interrupted unless there is very strong evidence that the student is thrashing in unprofitable ways. Second, the student is free to disregard the advice offered by the local agents. For this reason it is not critical that the local agents have a complete or fully accurate cognitive model of the student's misconception. If students believe the tutor has made an inappropriate suggestion, they can simply proceed with their plan. Finally, although local, rules embedded in the agents are still principled. While they do not represent a complete cognitive theory of expertise in inquiry, they are founded on a solid empirical base of observations of students using our microworlds, buttressed by commonsense. For example, although we do not yet possess a complete task analysis of scientific inquiry, most of us would agree that manipulating one variable at a time is good experiment planning.

Depending on the size of the islands of tutor control, an ILE that has been supplemented with locally intelligent agents begins to resemble a semi-intelligent tutoring system whose expert system is more global than local, but less than omniscient. Thus, as ILEs and ITS adapt to solve some of the central competence problems that plague them, we will see in the future the emergence of *mixed-initiative* tutors, blending student- and tutor-control of learning interactions in a wide variety of ways. The literature already shows evidence of this trend. We described White and Frederiksen's (1986) system for learning about electronic circuits as an ILE, but as they point out, progressions of qualitative models can be used in a variety of different learning regimes, ranging from very unstructured and student-controlled, to highly structured and tutor-controlled. Similarly, Fischer discusses active help systems (Fischer, Lemke and Schwab, 1985) and critiquing systems (Fischer, Lemke and McCall, 1990). Like ILEs in general, both active help and critiquing system are predicated on high student-control of learning. But, on occasion, they will use their expertise to unobtrusively volunteer information to learners and to critique suboptimal solutions. Elsom-Cook (1990) also describes a collection of guided discovery-tutoring systems that possess varying degrees of student- and tutor-initiative.

Expanding methods of teaching and goals for learning

In general, there are many creative approaches to developing flexible mixed-initiative systems. This area of research and system-building is relatively new, and it is probably premature to attempt to foresee all the diverse ways knowledge can be used in mixed-initiative intelligent systems for learning. Indeed it may be misleading to view mixed-initiative systems as differing just along a single continuum of student-tutor control. More generally, we expect a central future concern of AI applications will be to create mixed initiative systems that *define and implement a wide range of methods of learning and teaching*. Educational technology is now largely limited to the few methods of learning and teaching we have discussed in this paper. However, in addition to drill-and-practice and inquiry, there are many other ways to learn. Some new methods are just beginning to be examined in the educational technology literature. We list only a few below, to underscore their variety:

- *Case-based learning.* Schank and his colleagues (Schank and Edelson, 1990; Feifer and Soclof, 1991) have implemented several case-based learning systems in which students acquire expertise (e.g., in military history or telephone communication skills) by accessing cases from a rich library of past experiences. The cases selected must bear on a current situation to be resolved. Thus, instead of learning abstract rules to apply to situations, students exploit the analogies found in familiar cases to synthesize their own decision rules.
 - *Learning through coaxing and reflection.* How smart does a companion have to be to help you learn? ITS take the position that good one-on-one coaching demands deep knowledge of both the subject and student. Yet most of us have had the experience of improving the clarity of our ideas through discussions with friends who do not understand what we are doing but who can ask questions that cause us to restructure our thoughts. Kass (Kass, 1990; Kass and Guralnick, 1991) has recently developed systems that explore the efficacy of such coaxing and reflecting in learning.
 - *Simulation-based learning.* Perhaps the most effective training device yet developed is the flight simulator. It provides large amounts of inexpensive time-on-task to perfect skills that are too costly, slow, and dangerous to acquire using real aircraft. Building on this insight, researchers are now developing simulation-based learning environments for a wide range subjects from biology (Schank and Edelson, 1990) to social skills (Bevis and Kass, 1991).
 - *Visualization.* Tools for visualization exploit the computer's ability to render complex patterns succinctly in graphical representations (OSTP, 1992). Visualization often permits learners and professionals to get broad pictures of data, to "generate intuitions", and to suggest hypotheses for further testing. In some cases visualization may even substitute for symbolic reasoning and proof, not merely supplement it (Davis and Hersh, 1980). In these areas, then, visualization tools are providing new perceptual methods to more simply accomplish activities that, previously, required much subtle symbolic reasoning. But in other cases, visualization technologies are opening up completely new fields of study and defining new goals for curricula as well as providing new methods of learning. Chaos (Gleik, 1987) and fractals (Peitgen, Jurgens and Saupe, 1991) are two of the most dramatic examples of fields that simply did not exist before computer graphics provided the means to understand non-linear processes in a visual fashion.
 - *On-demand learning.* Fischer (1991) has developed prototype systems and a conceptual framework for learning on demand. He argues that this method of learning and teaching situates learning in a work context, instead of relegating it to a separate phase, and makes learning relevant to the task at hand. He also suggests that learning on demand or "just in time" learning will be increasingly important -- future workers will need to continually improve their skills.
 - *Apprenticeship.* Collins, Brown, and Newman's (1988) seminal paper has stimulated the development of new computer-based environments that provide some of the desirable features of traditional apprenticeship (see e.g., Lave, 1988) as a model of learning and teaching. Many of these new ideas and systems move apprenticeship mentoring activities -- modeling, coaching, and fading -- from training physical skills, like tailoring, to cognitive skills, like writing, reading, or mathematics.
 - *Collaboration.* Computer-based technologies for collaboration are providing new cooperative methods for work and learning. Applications of artificial intelligence in education typically support processes of individual learning, not just because these processes are effective, but because most research on learning in AI and cognitive science has focused on individual cognition. However, powerful network-based technologies are now making possible novel methods of group problem solving and learning. Networks permit cooperative work within classrooms, across classrooms (Scardamalia and Bereiter, 1993), and even across countries (Hunter, 1993). Collectively, they pose a serious challenge to the "solo" methods of learning and teaching institutionalized in today's classrooms.
- These examples demonstrate how AI, and information technologies in general, are beginning to expand the

available methods of learning and teaching in at least two ways. First, they may rejuvenate venerable learning and teaching methods like apprenticeship and one-on-one tutoring. Some of these methods are already relatively well understood and highly regarded but have not found their way into classrooms on a large scale, in part because they are labor-intensive. For example, although one-on-one tutoring leads to impressive student outcomes, at this time giving each student his or her own tutor would be prohibitively expensive. New technologies may provide the key to lowering the cost of implementing these familiar methods.

Second, the same technologies that breath life into traditional teaching techniques also can provide a foundation for completely new methods of acquiring knowledge. Visualization and collaboration, in particular, have been adjuncts to learning in the past, not primary methods of acquiring knowledge. For example, a diagram in geometry may orient the student in conducting a symbolic proof (or it may mislead her). But today proofs themselves may be visual. In other words, through new computer technologies visual images that once aided reasoning and learning are now becoming ways to reason and learn.

The transforming impact of new methods of learning and goals for learning

In the past, educators have looked to educational technology as a possible way to increase productivity of students and teachers in relatively straightforward ways; for example, to increase student learning rates, or to streamline teacher's routine practices. Past technologies, including most CAI systems, and most ITS, have implicitly offered education a set of tools to do better what they already do. They do not aim at new goals for student learning, nor do they expect outcomes to be achieved through new learning methods. In this respect, they imply relatively modest changes to the delivery of education.

Some future applications of educational technology -- even AI based applications -- may follow this familiar route. But we argue that most successful educational technology in the future must be part of a larger technology revolution. In the workplace, new information technologies, like visualization and collaboration tools are redefining how professionals do their jobs and what those jobs are. The needs of the workplace, in turn, are placing demands for new skills on education and training institutions. At the same

time, new information technologies provide novel means for meeting new workplace demands. Drill-and-practice and lecture methods of teaching have not institutionalized in classrooms because they are optimal methods of communicating information and acquiring knowledge. Such traditional classroom methods have survived primarily because they are reasonable vehicles for learning given the highly limited resources available to schools. As technologies in the workplace -- and our culture in general -- begin to reshape valued educational goals, they will also redefine the available resources for education. In the terms we have used above, they will permit us to consider a wide range of new methods of teaching and learning, and perhaps also to realize old methods -- like individualized tutoring -- that we have always believed to be valuable, but that hitherto have been too costly to implement on a wide scale. In summary, then, the long-term role of AI (and computer-based technology in general) in education will not be to support traditional teaching and learning practices, but to challenge and even threaten them by suggesting new things to learn and offering new ways to acquire them. For this reason it is misleading to view AI or technology in general as a means of *saving* education. At best, these forces will *transform* schools and classrooms, not improve them in any simple sense.

Moving new technologies into education

However, just as new methods of learning and teaching pose a threat to traditional classrooms, current educational practices pose many challenges to them. Many past technologies, including most CAI systems and ITS have implicitly offered education a set of tools to do better what they already do. They do not aim at new goals for student learning, nor do they expect outcomes to be achieved through new learning methods. In this respect, they imply relatively modest changes to the delivery of education. Yet, taken as a whole, even these technologies have not substantially improved educational outcomes. But new technologies -- including ILEs, mixed-initiative systems, as well as other approaches like visualization and collaborative tools -- do not promise to "fine tune" the practices of today's classroom; they offer new goals and practices for learning and teaching. Given the resistance of education to even relatively modest change, what hope do we have to move these new ideas into education on a broad scale?

Division of labor in implementing new learning goals and methods of teaching

Certainly moving new technologies into the classroom will require more than the scientific research and engineering required to develop potentially valuable educational applications and the funding necessary to make them available to all schools. We have noted that when using our ITS in classrooms 75% our effort was spent on technical and research issues, but implementation required most of our time when we integrated ILEs -- and the new goals and methods for learning they imply -- into schools. As technology continues to transform the goals for student learning and to enlarge the range of methods for teaching and learning, implementation will require proportionally more effort. In general, implementation tasks must develop:

- *New curricula.* Curricula will need to focus on new goals for learning and teaching, and must include materials and ideas for both teachers and students. These materials must help prepare both students and teachers for their roles in the new methods of teaching and learning we wish to institutionalize.
- *New methods and instruments for assessment.* If we change the goals of learning and teaching, it will be critical to evolve ways of measuring the new skills and outcomes we now value. As mentioned, today many important new learning outcomes remain relatively ill-defined. But if we wish these goals and outcomes to be widely accepted, it will be crucial that the instruments we provide be objective, reliable, valid, and cost-effective to apply.
- *New teaching practices and professional standards.* If teachers are to move from lecture and drill-and-practice methods of teaching to new methods like inquiry- or project-based learning, general standards will need to be established for these new methods, along with rich collections of specific teaching practices, strategies, and tactics.
- *Teacher education.* New goals and methods for teaching and learning imply extensive changes in teacher education and in the content and practices of schools of education. Today, most schools of education are barely equipped to provide teachers

with simple skills for presenting lectures and managing student drill-and-practice. But new teachers will require much more extensive support to successfully manage classes structured around new methods of learning like inquiry, or apprenticeship.

For the most part, researchers and educators developing applications of artificial intelligence and advanced technologies in education are involved in "proof of principle" studies. The intent behind prototype systems like our microworlds, for example, is to show that the appropriate use of new technologies can lead to valued educational outcomes. But as these systems begin to aim more at new methods of teaching and learning, and new goals for learning, we inevitably become involved in many of the complex activities enumerated above. To field our microworlds and new curricula for computer-based statistics (McArthur, Robyn, Lewis, and Bishay, 1992), for example, we developed new curricula and tests, and we provided extensive teacher training. Similarly, Harel (Harel and Papert, 1990) clearly spent more time understanding and shaping the roles of teachers and students, and evaluating what students had learned, than developing the LOGO tools themselves.

In the future, when working on a small scale it will remain feasible to develop new technology prototypes, define new goals and methods for learning, and develop new curricula and evaluation techniques, all as part of a single ambitious project. But we believe that scaling up these successes broadly will require a division of labor -- different groups or projects working in a coordinated fashion to put together the technology, curricula, assessment tools, professional standards, and teacher training pieces of a package of broad educational reform.

How tightly coupled do these different activities or projects need to be? We cannot answer this question definitively, but several general comments are worth making. A pragmatic optimist's approach suggests that eventually all the pieces will fit together properly without the benefits of policies that attempt to aggressively promote coordination. In educational technology discussions, this belief enjoys considerable popularity today. For example, Collins (1991) acknowledges the failure of past technologies, including film, radio, and TV, to penetrate education. But he argues that the new generation of computer technologies are fundamentally

different. He suggests that, because communication technologies are transforming work, and indeed the entire culture outside education, schools will eventually appropriate these tools, first for purposes they value, and later for goals that are becoming socially valuable. Even Cuban (1993) has partially recanted his pessimistic view of the potential of educational technology. This view does not deny that many distinct activities are necessary for successful reform. But it suggests that the pieces may be spawned and organized by a powerful "technology push" without planning or explicit coordination.

We are similarly optimistic about the potential value of new technologies for education. However, we believe that a tighter coordination between different groups of researchers and educators involved in developing technologies, evaluating them, designing new curricula, and providing teacher training, could have several benefits. In the past, such coordination has been rather loose. Researchers who develop education technology prototypes often have relatively little background in curriculum change and educational reform. Their work is usually funded by programs that are distinct from programs for teacher training and enhancement or curriculum development. They are usually more aware of what new technologies make possible in education than what education might like to have today. On the other hand, curriculum reform efforts, for example, often proceed with relatively little knowledge of how new technologies might change their goals and methods.

This relatively loose coordination has led to various missed opportunities. Inquiry learning, for example, has long been held in high regard by intellectual communities and technologists. But communities of teachers and administrators have operated largely independent of these views, perhaps in part because they have understood how chaotic these methods of teaching might be if imposed on traditional classroom cultures. More recently, NCTM (National Council of Teachers of Mathematics, 1991) has proposed new curriculum and professional development standards for mathematics that do make some significant changes to classroom practice and content. Still, a prevailing opinion of many educational technology researchers is that these new standards ignore opportunities for curriculum restructuring that new technologies uniquely afford.

A final example of missed opportunities comes from the our own research community, which spends most of its time developing "proof-of-principle" software prototypes. These projects have produced many exciting ideas for new educational tools, but almost all of these stop at very small-scale demonstrations. Only a few of these ideas are ever turned into products that find their way into classrooms; the rest are left on the shelf. Mechanisms like the National Diffusion Network (NDN) have been designed to act as a clearing-house for successful software, to encourage its dissemination to classrooms. But NDN by itself is sufficient only if the software fits easily into existing classrooms, and is consistent with traditional goals and methods of teaching. The problem, as we have repeatedly stressed, is that the most promising new ideas and software systems transform goals and methods of teaching. Simply providing a clearing-house for these tools will not solve the host of challenges that confront successful implementation of innovations of this magnitude.

The feature common to these missed opportunities is a failure of communication among the different stakeholders and role players in the development and deployment of educational technologies. In the future it will be important to consider new policies that improve the communication and coordination between these groups. A variety of policy options are worth examining, including larger consortium-based projects that combine software developers, teacher education institutions, and other key stakeholder groups; smaller separate projects that work together synergistically using new networking technologies; and incentives that bring high-tech companies into better cooperation with educational technology research and classroom practice. Policies to improve coordination of educational technology research, development, and deployment have always been important. Today they are essential. The opportunities for technology-driven change in education are more promising now than ever before. But the requirements for effective educational change -- to implement new goals for learning and new methods of learning and teaching rather than to fine-tune existing goals and methods -- are huge barriers that stand in the way of these promises.

Constructionism versus Instructionism and AI in Education

In the previous sections, we have discussed future mixed-initiative systems. We have viewed them from several perspectives, mentioning the technical challenges

confronting the development of intelligent and flexible systems, the need to operationalize new goals or outcomes for learning, and the challenges in implementing diverse new methods of teaching and learning on a broad scale. We close with a final perspective. The diverse mixed-initiative systems we see arising in the future also impact various theoretical debates about learning and knowledge. In this final section we briefly touch on one such debate.

Increasingly, we see in the literature strong theoretical distinctions between the ideas underpinning ITS and ILEs. On the one hand, a constructivist or "constructionist" view of learning and knowing (Papert and Harel, 1991) is used to argue for the student-centered style of pedagogy exemplified by ILEs. According to this view, knowledge must be built by the learner, piece by piece, and ILEs -- when supported by peers and mentors in a culture of learning -- are seen as ideal tools for empowering such self-guided construction. On the other hand, ITS are seen as justified by an "instructionist" approach to learning, and consistent with a "commodity" view of knowledge in which the tutor's goal is to transfer to the student some relatively static package of information about a topic. On this view, if we can analyze the knowledge we wish to transfer to the student into constituent parts, then the most efficient way to transfer this knowledge to students should be to carefully tutor it piece by piece. While these views of knowledge and learning may indeed be distinct, it is unclear to us whether they should be tightly associated with the distinction between ILEs and ITS, especially considering that the boundary between ILEs and ITS is becoming less precise.

In the heated debate between learning by construction versus instruction, it is easy to overlook that the disagreement has an empirical base. Specifically, the central disagreement concerns who is in the best position to make key decisions needed to promote learning. Is a knowledgeable teacher (automated or otherwise) essential to choose appropriate tasks, to provide rich and individualized feedback, and to situate learning in authentic tasks? Or can students -- working with tools that act as intelligent cognitive amplifiers and in a supportive context that includes peers and mentors -- make these decisions themselves? Can students in rich but passive environments generate better feedback for learning than a tutor might provide? Are students typically in a better

position to know what information they need for learning than a teacher?

To answer these questions, several research strategies are feasible. One pragmatic approach is to push the extreme alternatives as far as they will go. For ITS, this means seeing how many of the functions of human tutors can be usefully automated and assessing the quality of student learning they engender. Can we develop computer-based agents that are able to suggest interesting tasks at the edge of a students' skills? Can similar agents coach students and provide adequate feedback? For which subjects and topics of learning is it possible to develop such agents? And, what quality of learning do such highly tutor-controlled environments yield? Do they encourage a shallow procedural understanding of skills? Or can they help students learn a more valuable deeper understanding of concepts?

By the same token, pushing the ILE alternative as far as it will go means developing increasingly effective empowering tools and technologies for learning, embedding them in congenial learning environments, and demonstrating the efficacy of largely self-guided or inquiry-based learning. How dramatically can such ILEs magnify self-directed learning skills? How crucial for learning is the support of peers, mentors, and the broader culture of learning? And, what quality of learning do such highly student-controlled environments yield? Do students get stuck in conceptual "local maxima", and do they acquire ideas inefficiently because their searches too often wander? Or do they develop a deeper conceptual understanding of ideas and of inquiry itself?

Today, popular opinion in the educational research community (although certainly not in classroom practice) favors the constructive side of the debate. Constructivism enjoys this position *in spite of the fact that none of the empirical questions sketched above has been answered definitively in favor either side*. The best we can do now is guess at how future research will resolve the disagreements.

What are the likely answers to these questions? Recent research in human tutoring (Leinhardt, 1989; McArthur, Stasz and Zmuidzinas, 1990), and the increasing popularity of mixed-initiative tutoring systems both suggest that the best learning environments neither will be completely controlled by students nor by tutors. Most

effective learning environments include a mix of direct teaching, more passive support for learning, together with substantial student choice. For example, in our studies of inquiry-based tutoring by humans (Lewis, McArthur, Stasz, and Zmuidzinas, 1990; Lewis, Bishay, McArthur and Chou, 1992) even teachers experienced in orchestrating student-centered inquiries often interpolated bouts of lecture and tutor-directed coaching. Similarly, much work developing ILEs and mixed-initiative systems implicitly or explicitly recognizes the need for active support of learning (e.g., Brown and Duguid, 1993; Roschelle (in press). However, active roles in learning are typically associated with mentors, co-workers and peers, who engage learners in rich dialogues. In these environments, the role of technology is largely to provide a useful collection of tools that can amplify, enhance, or even transform the nature of these dialogues where learning takes place.

But as mixed-initiative systems begin to mature, there is no reason that technologies cannot take on more ambitious roles. An omniscient and highly-controlling ITS may be inappropriate model for the use of technology in many effective learning environments. However, as we have noted, this is no longer the only alternative to relatively passive ILEs. Instead, in the future we can expect to see mixed-initiative systems that include not only tools which amplify the students' inquiry but also locally intelligent agents which provide many of the active supports now supplied by peers and mentors. Today, for example, we encourage students learning statistics to ask members of their project team to critique experimental designs. Tomorrow, that team can include a local computer-based expert that has useful (but not complete) knowledge of experimental design.

Until we answer the empirical questions outlined above, it is unclear precisely which roles locally intelligent computer-based agents should play in learning, and how dominant they will be. Of course, to some extent the appropriate active tutoring role of such agents will depend on the subjects to be learned, on the background experience of students, and on their learning styles. If we believe it is still valuable for students to learn symbol manipulation skills in algebra, for example, an environment heavily populated with guiding agents -- approximating an ITS -- may be the most effective way to learn. If students are learning

inquiry skills themselves, or how to "brainstorm" in groups, then locally intelligent agents may play very modest roles. Similarly, a novice student may learn best in environments that include agents which can intensively model and coach formative skills (Collins and Brown, 1987). As the student acquires expertise these agents will probably "fade", permitting much more student initiative.

In some cases, then, locally intelligent agents may profitably play substantial teaching roles, but in many cases the roles of intelligent agents will be modest. However, we would like these roles to be modest because we believe that such roles will lead to the best student outcomes, not simply because we are technically unable to develop agents that are locally smart enough. We would like to limit the role of intelligent agents in education by principled choice, not by practical necessity. Today, we believe that most of what is wrong in applications of AI in education, including ITS, is that they are technically limited as models of human pedagogical expertise, not that they are wrong in principle.

Details aside, the main point is that ideas from artificial intelligence and knowledge-based systems neither support "instructionist" or "constructionist" views of teaching and learning wholeheartedly; rather, they can and will be used to implement a diverse set of methods of learning and teaching, perhaps aiming at different kinds of learning outcomes. Certainly, well-designed locally intelligent agents may on occasion strongly control a learning interaction -- much as Socrates did in his dialogue with the slave in Plato's *Meno*. But it is a naive caricature to assume that future applications of AI in education will be subject to the same limitations that befell first-generation ITS. Future mixed-initiative applications will not necessarily teach just through drill-and-practice or lecture. They will not "program" students to behave like a rigid procedure, nor will they necessarily assume that each task has only one "right answer". As AI expands to provide models of subtle reasoning skills, new systems will not be limited to tutoring routine procedural skills. Similarly, locally intelligent systems, like good human tutors, will learn to confront the challenges of teaching without "knowing everything" about the topics students learn. And although locally intelligent agents for learning will certainly transform the roles of teachers (and peers) in the classroom, they will not pretend to replace them.

References

- Anderson, J.R. (1983).** *The Architecture of Cognition*. Cambridge MA: Harvard University Press.
- Anderson, J.R., Boyle, D.F, and Reiser, B.J. (1985).** *Intelligent Tutoring Systems, Science Vol. 228, pp. 456-462.*
- Anderson, J.R., Boyle, D.F., Farrell, R, and Reiser, B.J. (1987).** *Cognitive principles in the design of design of computer tutors. In P. Morris (Ed.), Modeling Cognition, Wiley.*
- Anderson, J., Boyle, D .F., and Yost, G. (1985).** *The geometry tutor. Proceedings of the Ninth International Joint Conference on Artificial Intelligence.*
- Anderson, J. R., and Skwarecki, E. (1986).** *The automated tutoring of introductory computer programming, Communications of the ACM, Vol. 29, 9, 842-849.*
- Ausubel, D. P. (1961).** *Learning by discovery: Rationale and mystique. Association of Secondary School Principals, 45, 18-58.*
- Berman, S. and McLaughlin, M. (1978).** *Federal Programs Supporting Educational Change: Vol. VIII, Implementing and Sustaining Innovations, The RAND Corporation, R-1589/8-HEW.*
- Bevis, E, and Kass, A. (1991).** *Teaching by Means of Social Simulation. In Proceedings of the International Conference on the Learning Sciences, Evanston, pp. 45-51.*
- Bloom, B. S. (1984).** *The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring, Educational Researcher, 13, 6, June/July.*
- Bonar, J. (1991).** *Interface architectures for intelligent tutoring systems. In H. Burns, J. Parlett, & C. Luckhardt Redfield (Eds.), Intelligent Tutoring Systems: Evolutions in Design. Hillsdale NJ: Lawrence Erlbaum.*
- Brown, J. S., and Burton, R. R. (1978).** *Diagnostic models for procedural bugs in basic mathematical skills. Cognitive Science, 2, 155-192.*
- Brown, J. S., Burton, R. R., and de Kleer, J. (1982).** *Pedagogical, natural language and knowledge engineering and pedagogical techniques in SOPHIE I, II, and III. In D. H. Sleeman & J. S. Brown (Eds.), Intelligent Tutoring Systems (pp. 227-282). New York: Academic Press.*
- Brown, J.S. and Duguid, P. (1993).** *Stolen knowledge. Educational Technology, March, 1993.*
- Brown, S. I. and Walter, M. I. (1990).** *Problem Posing. Lawrence Erlbaum: Hillsdale NJ.*
- Bruneau, J. Chambreuil, A., Chambreuil, M, Chanier, M, Dulin, P, Lotin and Nehemie, P. (1991).** *Cognitive science, artificial intelligence, new technologies: How to cooperate for a computer-assisted learning to read system. In Proceedings of the International Conference on the Learning Sciences, Evanston, Il.*
- Bruner, J. S. (1961).** *The act of discovery. Harvard Educational Review, 31 (1), 21-32.*
- Burns, B., Gray, W. D., Radlinkski, E. R. (1991).** *Tuning the ideal student model: Towards and intelligent editing of ITS models. In Proceedings of the International Conference on the Learning Sciences, Evanston, Il.*
- Burton, R. R., and Brown, J. S. (1982).** *An investigation of computer coaching. In D. H. Sleeman & J. S. Brown (Eds.), Intelligent Tutoring Systems (pp. 79-98). New York: Academic Press.*
- Clancey, W.J. (1987).** *Knowledge-based tutoring: The GUIDON program. Cambridge, MA: The MIT Press.*
- Cohen, D. K. (1988).** *Educational technology and school organization. In R. S. Nickerson & P. P. Zodiates (Eds.), Technology in Education: Looking Toward 2020 (pp. 231-264). Hillsdale, NJ: Erlbaum.*
- Collins, A. (1991).** *The role of technology in restructuring schools. Phi Delta Kappan, September, 28-36.*
- Collins, A., and Brown J.S. (1987).** *The computer as a tool for learning through reflection. In H. Mandl and A. Lesgold (Eds.), Learning issues for intelligent tutoring systems. New York: Springer-Verlag.*
- Collins, A., J. S. Brown, and S. E. Newman (1988).** *Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics. In L.B. Resnick (Ed.), Cognition and Instruction: Issues and Agendas. Hillsdale NJ: Lawrence Erlbaum.*
- Cooper, E. W. (1991).** *An architecture for apprenticeship: Collaboration with an intelligent tutoring system for qualitative electrical troubleshooting. In Proceedings of the International Conference on the Learning Sciences, Evanston, Il.*
- Cuban, L. (1986).** *Teachers and machines. New York: Teachers College Press.*
- Cuban, L. (1993).** *Computers meet the classroom; Classroom wins. Education Week, November 11, pp 26-27.*
- Davis, R. (1991).** *Constructivist views on the teaching and learning of mathematics, Journal for Research in Mathematics Education, Monograph 4, National Council of Teachers of Mathematics.*

- deKleer, J. and Brown, J. S. (1984).** *A physics based on confluences. Artificial Intelligence, 24, pp 7-83.*
- Du, Z., and McCalla, G. (1991).** *CBMIP -- A case-based mathematics instructional planner. In Proceedings of the International Conference on the Learning Sciences, Evanston, IL.*
- Elsom-Cook, M., (ed.) (1990).** *Guided Discovery Tutoring: A Framework for ICAI Research. London: Paul Chapman Publishing.*
- Feifer, R. G. (1989).** *An intelligent tutoring system approach to teaching people how to learn. Proceedings of the Eleventh Annual Conference of the Cognitive Science Society, Ann Arbor, Michigan.*
- Feifer, R. and Soclof, M. (1991).** *Knowledge-based tutoring systems: Changing the focus from learner modeling to teaching. In Proceedings of the International Conference on the Learning Sciences, Evanston, IL.*
- Fischer, G. (1991).** *Supporting Learning on Demand with Design Environments. Proceedings of the International Conference on the Learning Sciences, Evanston, pp. 165-171.*
- Fischer, G., Lemke, A., and Schwab, T. (1985).** *Knowledge-based Help Systems. Human Factors in Computing Systems, CHI 1985 Conference Proceedings, San Francisco, pp. 161-167, New York: ACM.*
- Fischer, G., Lemke, A., and McCall, R. (1990).** *Towards a system architecture supporting contextualized learning. Proceedings of AAAI-90, pp. 420-425, Cambridge MA: AAAI Press/MIT Press.*
- Foss, C. (1987).** *Acquisition of error management skills. Third International Conference on Artificial Intelligence and Education, pp. 27.*
- Forbus, K. (1984).** *Qualitative process theory. Artificial Intelligence, 24, pp 85-168.*
- Forbus, K. (1991).** *Towards tutor compilers: Self-explanatory simulations as an enabling technology. In Proceedings of the International Conference on the Learning Sciences, Evanston, IL.*
- Frederiksen, C., Donin, J., DeCary, M., and Edmond, B. (1991).** *Discourse-based second-language learning environments. In Proceedings of the International Conference on the Learning Sciences, Evanston, IL.*
- Frederiksen, J., White, B, Collins, A. Eggan, G. (1988).** *Intelligent Tutoring systems for electronic troubleshooting. In J. Psofka, D. Massey, and S. Mutter. Intelligent Tutoring Systems: Lessons Learned. Hillsdale NJ: Lawrence Erlbaum.*
- Geertz, C. (1973).** *The Interpretation of Cultures. New York: Basic Books.*
- Gleik, J. (1987).** *Chaos: Making a new science. New York: Penguin Books.*
- Harel, I. and Papert, S. (1990).** *Software Design as a Learning Environment Interactive Learning Environments, 1(1), 1-33.*
- Hunter, B. (1993).** *Internetworking: Coordinating technology for systemic reform. Communications of the ACM, 36(5), pp. 42-46.*
- Kass, A. (1990).** *Are electronic tutors really what we want to build? American Association for Artificial Intelligence Spring Symposium on Knowledge-based Environments for Learning and Teaching, Stanford University.*
- Kass, A. and Guralnick, D. (1991).** *Environments for Incidental Learning: Taking Road Trips Instead of Memorizing State Capitals. In Proceedings of the International Conference on the Learning Sciences, Evanston, p. 258-264.*
- Lakatos, I. (1976)** *Proofs and Refutations. New York: Cambridge University Press.*
- Larkin, J., H., McDermott, J., Simon, D. P., and Simon, H. A. (1980).** *Expert and novice performance in solving physics problems. Science, 208 1335-1342.*
- Lave, J. (1988).** *Cognition in Practice, Cambridge, Cambridge University Press.*
- Lawrence, E. (1970).** *The Origins and Growth of Modern Education. Baltimore, MD: Penguin Books.*
- Leinhardt, G. (1989).** *Math lessons: A contrast of novice and expert competence. Journal for Research in Mathematics Education, 20(1), 52-75.*
- Lewis, M, Bishay, M., McArthur, D. (1993).** *Supporting Discovery Learning in Mathematics: Design and Analysis of an Exploration Environment and Inquiry Activities. Submitted to Instructional Science.*
- Lewis, M., Bishay, M., and McArthur, D. (1993b).** *The Macrostructure and Microstructure of Inquiry Activities: Evidence from Students using a Microworld for Mathematical Discovery. Proceedings of the World Conference on Artificial Intelligence and Education, Edinburgh, August.*
- Lewis, M., McArthur, D., Bishay, M., and Chou, J. (1992).** *Object-Oriented Microworlds for Learning Mathematics through Inquiry: Preliminary Results and Directions. Proceedings of the East-West Conference on Emerging Computer Technologies in Education, Moscow, April.*

- Lewis, M. W., McArthur, D., Stasz, C., and Zmuidzinis, M. (1990).** *Discovery-based Tutoring in mathematics*. Paper presented at the American Association for Artificial Intelligence Spring Symposium on Knowledge-based Environments for Learning and Teaching, Stanford University, March.
- Lesgold, A. M., Lajoie, S. P., Bunzo, M., and Eggan, G. (1993).** *SHERLOCK: A coached practice environment for an electronics troubleshooting job*. In J. Larkin, R. Chabay, & C. Scheffic (Eds.), *Computer assisted instruction and intelligent tutoring systems: Establishing communication and collaboration*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lesgold, A., Eggan, G., Katz, S., and Rao, G. (in press).** *Possibilities for Assessment Using Computer-Based Apprenticeship Environments*. To appear in W. Regian and V. Shute (Eds.), *Cognitive approaches to automated instruction*. Hillsdale, NJ: Erlbaum.
- Lester, J. C. and Porter, B. W. (1991).** *A student-sensitive discourse generator*. In *Proceedings of the International Conference on the Learning Sciences, Evanston, IL*.
- London, B., and Clancey, W. J. (1982).** *Plan recognition strategies in student modeling: Prediction and description*. In *Proceedings of the 1982 National Conference on Artificial Intelligence*, [pp. 335-338].
- Matz, M. (1982).** *Towards a process model for high school algebra errors*, In D. H. Sleeman & J. S. Brown (Eds.), *Intelligent Tutoring Systems* New York: Academic Press.
- McArthur, D. and Lewis, M. (1991).** *Overview of object-oriented microworlds for learning mathematics through inquiry*. N-3242-NSF, RAND Corporation, Santa Monica, CA.
- McArthur, D. and Lewis, M. (1991b).** *Overview of object-oriented microworlds for learning mathematics through inquiry*. *Proceedings of the International Conference on the Learning Sciences, Evanston, IL*.
- McArthur, D., Lewis, M. W., Ormseth, T., Robyn, A., Stasz, C., and Voreck, D. (1989).** *Algebraic thinking tools: Supports for modeling situations and solving problems in kids' worlds*. *Technology and Learning*, 3(2).
- McArthur, D., Robyn, A., Lewis, M. and Bishay, M. (1992).** *Designing new curricula for mathematics: A case-study of computer-based statistics in high school*. RAND WD-5930-ED.
- McArthur, D., and Stasz, C. (1990).** *An intelligent tutor for basic algebra*. R-3811-NSF, RAND Corporation, Santa Monica, CA.
- McArthur, D., Stasz, C., Hotta, J., Peter, O., and Burdorf, C. (1988).** *Skill-oriented task sequencing in an intelligent tutor for basic algebra*. *Instructional Science*, 17, 281-307.
- McArthur, D., Stasz, C., and Zmuidzinis, M. (1990).** *Tutoring techniques in algebra*. *Cognition and Instruction*, 7(3), 197-244.
- Merrill, D. C., Reiser, B. J., Ranney, M., and Tafton, J. G. (1992).** *Effective tutoring techniques: A comparison of human tutors and intelligent tutoring systems*. *The Journal of the Learning Sciences*, 2(3), 277-306
- National Council of Teachers of Mathematics. (1989).** *Curriculum and Evaluation Standards for School Mathematics*. Reston, VA: NCTM.
- National Council of Teachers of Mathematics. (1991).** *Professional Standards for Teaching Mathematics*. Reston, VA: NCTM.
- Newell, A., and Simon, H.A. (1972).** *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Ohlsson, S. (1986).** *Some principles of intelligent tutoring*, *Instructional Science*, 14, 293-326.
- Ohlsson, S. (1991).** *System hacking meets learning theory: Reflections on the goals and standards of research in artificial intelligence and education*. *Journal of Artificial Intelligence and Education*, 2(3), 5-18.
- OSTP (1992).** *Grand Challenges: High Performance Computing and Communications*. The Committee on Physical, Mathematical, and Engineering Sciences, Federal Coordinating Council for Science, Engineering, and Technology. Executive Office of the President, Washington, D.C.
- Papert, S. (1980).** *Mindstorms: Children, Computers and Powerful Ideas*. Basic Books: New York.
- Papert, S. (1993).** *The Childrens' Machine*. New York: Basic Books.
- Papert, S. and Harel I. (1991).** *Constructionism*, Norwood NJ: Ablex Publishing.
- Pea, R. (1987).** *Cognitive technologies for mathematics education*. In Schoenfeld, A.H (Ed.), *Cognitive Science and Mathematics Education*, Hillsdale, NJ: Lawrence Erlbaum.
- Peitgen, H., Jurgens, H, and D. Saupe (1991).** *Fractals for the Classroom*. New York: Springer-Verlag.
- Polya, G. (1962).** *Mathematical Discovery*. New York: John Wiley and Sons.

- Psojka, J, Massey, D., and Mutter, S. (1988).** *Intelligent Tutoring Systems: Lessons Learned.* Hillsdale NJ: Lawrence Erlbaum.
- Putnam, R.T. (1987).** *Structuring and adjusting content for students: A study of live and simulated tutoring of addition,* *American Educational Research Journal, 24 (1), pp. 13-48.*
- Raghavan, K., Schultz, J., Glaser, R., and Schauble, L. (1989).** *A computer coach for inquiry skills.* Pittsburgh, PA: Learning Research and Development Center, University of Pittsburgh.
- Resnick, L. B. (1987).** *Education and Learning to Think.* Washington D. C.: National Academy Press.
- Resnick, M. (1991).** *Overcoming the centralized mindset: Towards an understanding of emergent phenomena.* In S. Papert and I. Harel (Eds.), *Constructionism,* Norwood NJ: Ablex Publishing.
- Robyn, A., Stasz, C., Ormseth, T., and McArthur, D. (1989).** *Implementing computer-assisted instruction in first-year algebra classes.* Paper presented at the American Education Research Association National Conference, San Francisco, April.
- Robyn, A., Stasz, C., McArthur, D., Ormseth, T., and Lewis, M. W. (1992).** *Implementing a novel computer-related algebra course.* RAND N-3326-NSF/RC.
- Roschelle, J. (in press).** *Learning by collaboration: Convergent conceptual change.* *Journal of the Learning Sciences.*
- Scardamalia, M, and Bereiter, C. (1993).** *Technologies for knowledge-building discourse.* *Communications of the ACM, 36(5), pp. 37-41.*
- Schank, R. and Edelson, D. (1990).** *A role for AI in education: Using technology to reshape education.* *Journal of Artificial Intelligence in Education, 1(2), 3-20.*
- Schoenfeld, A.H. (1985).** *Mathematical Problem Solving.* New York: Academic Press.
- Schofield, J. W., Evans-Rhodes, D, and Huber, B. R. (1990).** *Artificial Intelligence in the Classroom:* *The Impact of a Computer-based Tutor on Teachers and Students,* *Social Science Computer Review, 8(1), 24-41.*
- Shute, V., and Bonar, J. (1988).** *Intelligent tutoring systems for scientific inquiry skills.* Unpublished manuscript LRDC.
- Shute, V., and Glaser, R. (1990).** *Large-scale evaluation of an intelligent discovery world: SMITHTOWN.* *Interactive Learning Environments, 1, 51-77.*
- Shute, V., Glaser, R., and Raghavan, K. (1988).** *Inference and discovery in an experimental laboratory.* Tech. Rep. LRDC.
- Sleeman, D.H., and Smith M.J. (1981).** *Modeling Student's Problem Solving,* *Artificial Intelligence, 16, pp. 171-188.*
- Sleeman, D. and Brown, J. S. (1982).** *Intelligent Tutoring Systems* New York: Academic Press.
- Steen, L. A. (1988).** *The science of patterns.* *Science, 240, 611-616.*
- Steen, L. A. (1990).** *On the Shoulders of Giants,* Washington DC: National Academy Press.
- Swanson, J. (1990).** *The effectiveness of tutorial strategies: An experimental evaluation.* Paper presented at the annual meeting of the American Educational Research Association, Boston MA, April.
- Wenger, E. (1987).** *Artificial Intelligence and Tutoring Systems.* Los Altos CA: Morgan and Kaufmann.
- White, B.Y., and Frederiksen, J.R. (1986).** *Qualitative models and intelligent learning environments.* In Lawler, R., & Yazdani, M. (Eds.) (1987). *Artificial Intelligence and Education: Learning Environments and Intelligent Tutoring Systems.* Norwood NJ: Ablex.
- Woolf, B. (1991).** *Representing, acquiring, and reasoning about tutoring knowledge.* In H. Burns, J. Parlett, & C. Luckhardt Redfield (Eds.), *Intelligent Tutoring Systems: Evolutions in Design.* Hillsdale NJ: Lawrence Erlbaum.
- Zuboff, S. (1988).** *In the Age of the Smart Machine: The Future of Work and Power,* Basic Books.

